

AD A103562

~~LEVEL~~

16567.10-1

(1)

CENTER FOR PURE AND APPLIED MATHEMATICS
UNIVERSITY OF CALIFORNIA, BERKELEY

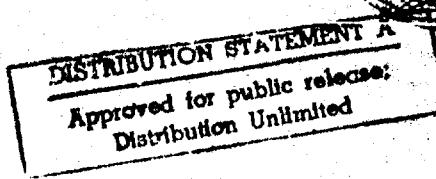
PAM-49

HORSESHOES AND ARNOLD DIFFUSION FOR HAMILTONIAN
SYSTEMS ON LIE GROUPS

Philip J. Holmes and Jerrold E. Marsden



FILE COPY



81 9 01 093

July 28, 1981

Legal Notice

This report was prepared as an account of work sponsored by the Center for Pure and Applied Mathematics. Neither the Center nor the Department of Mathematics, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information or process disclosed.

UNCLASSIFIED

(18) HRC/1

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM	
1. REPORT NUMBER (P) 16567.10-M	2. GOVT ACCESSION NO. AD-A103 562	3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (or SUBTITLE) (6) Horseshoes and Arnold Diffusion for Hamiltonian Systems on Lie Groups	5. TYPE OF REPORT & PERIOD COVERED (7) Technical 10pt.5		
7. AUTHOR(s) (10) Philip J. Holmes Jerrold E. Marsden	8. CONTRACT OR GRANT NUMBER(s) (15) DAAG29-79-C-0086 VNSF-ENG78-02891		
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of California Berkeley, CA 94720	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS		
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709	12. REPORT DATE (11) 28 Jul 81		
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (12) 62	13. NUMBER OF PAGES 59		
15. SECURITY CLASS. (of this report) Unclassified			
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited. (14) PHM1-49, TR-16			
17. DISTRIBUTION STATEMENT (of the material enclosed in Block 30, if different from Report) NA			
18. SUPPLEMENTARY NOTES The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.			
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)			
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The Melnikov theory of perturbations of Hamiltonian systems containing homoclinic orbits is extended to systems containing canonical variables belonging to the coadjoint orbits of a Lie group. This is applied to the free rigid body with attachments and to the nearly symmetric top. These systems are thereby shown to have transverse homoclinic manifolds in an appropriate return map and therefore have complex dynamics. In particular, the heavy top and rigid body with one attachment are shown to contain horseshoes and therefore have no additional analytic integrals, while the rigid body with several			

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 85 IS OBSOLETE

304950

attachments exhibits Arnold diffusion.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

1

Technical Report #16

P.J. Holmes and J.E. Marsden

July, 1981

DTIC
SELECTED
SEP 1 1981
S H

U.S. Army Research Office
Contract DAA G29-79-C-0086

University of California, Berkeley

Approved for public release;
distribution unlimited.

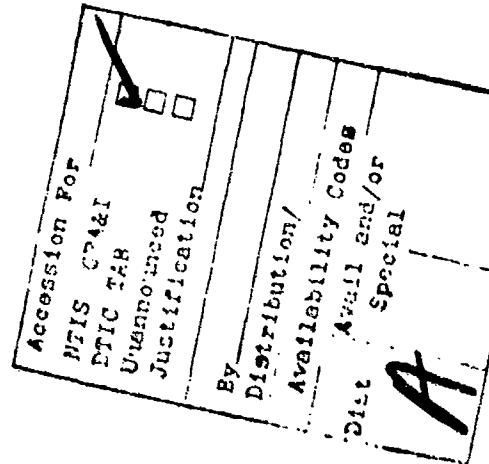
DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

Horseshoes and Arnold Diffusion for Hamiltonian
Systems on Lie Groups

Philip J. Holmes* and Jerrold E. Marsden**

1. Introduction
2. Hamiltonian systems on Lie groups and semi-direct products
3. The rigid body
4. Transversal intersections of invariant manifolds for KKS variables coupled to action angle variables
5. An example: the rigid body with attachments
6. Melnikov theory for systems with S^1 symmetry
7. An example: the motion of a nearly symmetric heavy top.
8. Conclusions

References



Research partially supported by ARO Contract DAAG-29-79-C-0086 and by
NSF Grants ENG 78-02891 and MCS-06718

* Department of Theoretical and Applied Mechanics, Cornell University,
Ithaca, New York, 14853.

** Department of Mathematics, University of California, Berkeley,
California, 94720

Abstract

The Melnikov theory of perturbations of Hamiltonian systems containing homoclinic orbits is extended to systems containing canonical variables belonging to the coadjoint orbits of a Lie group. This is applied to the free rigid body with attachments and to the nearly symmetric top. These systems are thereby shown to have transverse homoclinic manifolds in an appropriate return map and therefore have complex dynamics. In particular, the heavy top and rigid body with one attachment are shown to contain horseshoes and therefore have no additional analytic integrals, while the rigid body with several attachments exhibits Arnold diffusion.



§1. Introduction

This paper presents theorems which establish the existence of horseshoes and Arnold diffusion for nearly integrable Hamiltonian systems associated with Lie groups. The methods are based on our two previous papers, Holmes and Marsden [1981b,c]. The two main examples treated here are as follows:

1. The rigid body with attachments. This system has horseshoes (with one attachment) and Arnold diffusion (with two or more attachments).
2. A rigid body under gravity, close to a symmetric (Lagrange) top.

This system is shown to have horseshoes (and hence is not integrable).

The main new feature present here is the presence of Lie groups. Both the symmetry groups and the basic phase spaces involve Lie groups and our perturbation methods must be modified to take this into account. As in our previous work, the results hinge on reduction together with a method of Melnikov. This is used to analyze the perturbation of a homoclinic orbit in an integrable Hamiltonian system. In example one, the unperturbed system is the free rigid body which has a homoclinic orbit lying on a sphere. This sphere arises as the coadjoint orbit for the rotation group $SO(3)$ and the computation of Poisson brackets needed in the Melnikov theory is most easily done using the KKS (Kostant, Kirillov-Souriau) theory of coadjoint orbits. This theory is reviewed in Section 2.

After removing the angular momentum about the vertical axis, the correct phase space for a rigid body under gravity is T^*S^2 the cotangent bundle of a sphere. This and its connection with Euler angles and coadjoint orbits in the Euclidean group is explained in Section 3. This section thus sets

up the basic phase spaces needed in the analysis of our second example.

Section 4 develops the Melnikov theory when the phase space is a product of a KKS phase space and a set of action angle variables. This is applied to the rigid body with attachments in Section 5.

Section 6 develops the Melnikov theory for systems on a phase space where the unperturbed system admits an S^1 symmetry and has a homoclinic orbit in the reduced phase space. This generalization does not assume the original phase space is simply a product. This generalization is needed for and is applied to the nearly symmetric heavy top in Section 7.

The two examples presented here were selected because of their physical interest and because they provide good models for how the Melnikov theory must be modified for systems with topologically nontrivial phase spaces.

The presence of horseshoes in the motion of a nearly symmetry heavy top implies, amongst other things, that the dynamics is complex and cannot be captured by averaging methods (cf. Akulenko, Leshchenko and Chernousko [1979]), that the dynamics has periodic orbits of arbitrarily high period embedded in an invariant Cantor set and that the system admits no additional analytic integrals. The latter fact is consistent with known classical results, but the existence of horseshoes is a stronger and, we think, more significant assertion. Ziglin [1981] recently proved a general nonintegrability theorem that includes the non-standard rigid body cases but his proof proceeds along different lines and does not seem to yield horseshoes.

We expect that dissipative and forcing terms added to these systems can be dealt with along the lines of Holmes and Marsden [1981b].

Acknowledgements. We thank Allan Kaufman, Steve Smale and Alan Weinstein for motivation and several useful comments.

§2. Hamiltonian systems on Lie groups and semi-direct products

Since the basic paper of Arnold [1966], Lie groups have played an important role in the construction of phase spaces and the symmetry properties of some important mechanical systems. For systems such as the rigid body one wishes to realize the classical Euler equations as Hamiltonian equations on an appropriate phase space to apply Hamiltonian perturbation techniques. For a rigid body free to rotate about its center of mass, the basic phase space is $T^*SO(3)$ which is conveniently parametrized by the Euler angles (ϕ, ψ, θ) and the corresponding conjugate momenta $(p_\phi, p_\psi, p_\theta)$. However the $SO(3)$ symmetry of the problem enables one to reduce this 3 degree of freedom problem to a one degree of freedom system whose phase space is a sphere in body-angular momentum space.

For purposes of this paper we are interested mostly in the Poisson bracket structure since it is this which directly enters the Melnikov theory. In what follows we describe the Poisson structure associated with a Lie group and its connection with coadjoint orbits. (For the same theory with emphasis on the symplectic geometry, see Abraham and Marsden [1978, Chapter 4].) Proofs may be readily supplied by referring to the preceding reference, to Arnold [1978] and to Guillemin and Sternberg [1980].

(a) Generalities on KKS structures

Let G be a Lie group and \mathfrak{g} its Lie algebra. For $\xi, \eta \in \mathfrak{g}$, $[\xi, \eta]$ denotes the Lie bracket of ξ and η . Let \mathfrak{g}^* denote the dual space of \mathfrak{g} . For $F: \mathfrak{g}^* \rightarrow \mathbb{R}$, and the variable in \mathfrak{g}^* denoted by μ , define $\frac{\delta F}{\delta \mu}: \mathfrak{g}^* \rightarrow \mathfrak{g}$ by

$$DF(\mu) \cdot v = \left\langle v, \frac{\delta F}{\delta \mu} \right\rangle, \quad (2.1)$$

where \langle , \rangle denotes the pairing between \mathfrak{g}^* and \mathfrak{g} ; and $DF(\mu) : \mathfrak{g}^* \rightarrow \text{IR}$ is the usual (Frechet) derivative. It is understood that $\frac{\delta F}{\delta \mu}$ is evaluated at the point μ .

The KKS-bracket of two functions $F, G : \mathfrak{g}^* \rightarrow \text{IR}$ is defined by

$$\{\{F, G\}\}(\mu) = -\mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta G}{\delta \mu} \right]. \quad (2.2)$$

This bracket makes the smooth functions from \mathfrak{g}^* to IR into a Lie algebra. The only non-obvious condition is Jacobi's identity.

Next we describe the relationship between the KKS bracket and coadjoint orbits. For $g \in G$, let $\text{Ad}_g : \mathfrak{g} \rightarrow \mathfrak{g}$ be the adjoint representation (the linearization of the map $I_g : h \mapsto ghg^{-1}$ at $h = \text{identity}$) and $\text{Ad}_{g^{-1}}^* : \mathfrak{g}^* \rightarrow \mathfrak{g}^*$ the coadjoint representation. For $\mu_0 \in \mathfrak{g}^*$, let $\mathcal{O} = \{\text{Ad}_{g^{-1}}^* \mu_0 \mid g \in G\}$ be the orbit of μ_0 . A theorem of Kirillov, Kostant and Souriau states that \mathcal{O} is a symplectic manifold. We now describe the symplectic structure. For $\mu \in \mathcal{O}$, tangent vectors to \mathcal{O} and μ have the following form: let $\xi \in \mathfrak{g}$ and define $\tilde{\xi}(\mu) \in \mathfrak{g}^*$ by $n \mapsto \langle \mu, [\xi, n] \rangle$. Then $\tilde{\xi}(\mu)$ is tangent to \mathcal{O} at μ . The formula

$$\omega_\mu(\tilde{\xi}(\mu), \tilde{\eta}(\mu)) = -\langle \mu, [\xi, \eta] \rangle \quad (2.3)$$

defines a symplectic form on \mathcal{O} . Formulas (2.2) and (2.3) are related as follows: for $F, G : \mathfrak{g}^* \rightarrow \text{IR}$,

$$\{\{F, G\}\}|_0 = \{F|_0, G|_0\}_0, \quad (2.4)$$

where $\{\ \}_0$ is the Poisson bracket computed from the symplectic structure (2.3) on \mathcal{O} .

If $H : \mathfrak{g}^* \rightarrow \text{IR}$ is a given Hamiltonian, there is a unique vector field x_H on \mathfrak{g}^* such that any function $F : \mathfrak{g}^* \rightarrow \text{IR}$ which evolves along the flow

of X_H satisfies

$$\dot{F} = \{\{F, H\}\} . \quad (2.5)$$

Explicit equations of motion for $\mu \in \mathfrak{o}_J^*$ can be computed directly from (2.5) and (2.2) in examples, without the computation of coadjoint orbits, by letting F be coordinate functions on \mathfrak{o}_J^* . The equations so obtained are thus

$$\dot{\mu} = X_H(\mu) . \quad (2.6)$$

The vector field X_H is necessarily tangent to the coadjoint orbits; thus if $\mu(0) \in \mathcal{O}$, then $\mu(t) \in \mathcal{O}$ as well. Furthermore, $X_H|_0 = X_{H|_0}$ where $X_{H|_0}$ is the Hamiltonian vector field on \mathcal{O} computed from H using the symplectic structure (2.3).

As we shall see in the next sections, the classical Euler equations for a rigid body (with or without gravity) can be expressed in the form (2.5).

Remark. In Arnold [1966] and Ebin and Marsden [1970] it is shown that the equations of an incompressible fluid also fall into this class using the group D_{vol} of volume preserving diffeomorphisms of space. For compressible flow, the appropriate group is the semi-direct product of diffeomorphisms and functions. (The reason semi-direct products are relevant is explained abstractly below). When appropriately coupled to the electromagnetic field, equations (2.5) also include the equations of plasma physics and magnetohydrodynamics (see Marsden and Weinstein [1981]).

For the rigid body free to rotate about a fixed point, the basic phase space one starts with is $T^*SO(3)$, irrespective of whether or not gravity is present. In the absence of gravity, reduction by $SO(3)$ leads naturally to

the KKS Poisson structure on the Lie algebra of $SO(3)$. However when gravity is present, reduction by S^1 leads to the KKS Poisson structure on the Lie algebra of the Euclidean group; i.e. the semi-direct product $SO(3) \in \mathbb{R}^3$. Proofs of assertions made in the following may be found in one or more of Abraham and Marsden [1978], Guillemin and Sternberg [1980] and Ratiu and van Moerbeke [1981].

(b) KKS Structures for Reduction by a Subgroup⁺

Let G be a Lie group and T^*G its cotangent bundle. Let $v \in \mathfrak{g}^*$ and let $G_v = \{g \in G \mid \text{Ad}_{g^{-1}}^* v = v\}$ be the isotropy subgroup of v . Now G_v acts on G by left translation and hence on T^*G . This action has an Ad^* -equivariant momentum map

$$j^v : T^*G \rightarrow \mathfrak{g}_v^*$$

where \mathfrak{g}_v is the Lie algebra of G_v . In fact, for $a_g \in T_g^*G$ and $\xi \in \mathfrak{g}_v$,

$$(j^v(a_g), \xi) = (a_g, T_{e^g}^* \cdot \xi) \quad (2.7)$$

where $R_g : G \rightarrow G$ denotes right translation by g and (\cdot, \cdot) denotes the pairings between \mathfrak{g}_v^* and \mathfrak{g}_v or T_g^*G and T_gG as appropriate.

Now let $\mu \in \mathfrak{g}^*$ and $\bar{\nu} \in \mathfrak{g}_v^*$ be its restriction to \mathfrak{g}_v . We shall henceforth make the (generic) assumption that G_v is abelian. Then the reduced space of T^*G relative to the action of G_v is

⁺The heavy rigid top is done two separate ways in Section 7, namely in terms of Euler angles and in terms of the KKS structure. Those who wish only to read the Euler angle proof may omit the rest of this section and the third part (c) of the next section.

$$p_{\mu, v} = (J^v)^{-1}(\bar{\mu})/G_v \quad (2.8)$$

Let $\langle\langle \cdot, \cdot \rangle\rangle$ be a left invariant metric on G and $K : T^*G \rightarrow \mathbb{R}$ the corresponding kinetic energy function: $K(\alpha_g) = \frac{1}{2} \langle\langle \alpha_g, \alpha_g \rangle\rangle$. Let α_μ be the one form on G defined at $g \in G$ by minimizing K over the affine space of α_g such that $J^v(\alpha_g) = \bar{\mu}$.

Remark. If $G_v = S^1$ and $\xi \in \mathfrak{o}_v$ is the element corresponding to $\bar{\mu}$ via $\langle\langle \cdot, \cdot \rangle\rangle$ i.e. $\langle\langle \xi, \omega \rangle\rangle = \bar{\mu}(n)$ for all $n \in \mathfrak{o}_v$, then for $v \in T_g G$,

$$\alpha_\mu(g) \cdot v = \langle\langle T_e R_g \xi, v \rangle\rangle \cdot \frac{\langle\langle \xi, \xi \rangle\rangle}{\langle\langle \text{Ad}_{g^{-1}} \xi, \text{Ad}_{g^{-1}} \xi \rangle\rangle} \quad (2.9)$$

The function $K(\alpha_\mu(g)) = v_\mu(g)$ is called the amended potential.

The map $\alpha_q \mapsto \alpha_q - \alpha_\mu$ induces a symplectic diffeomorphism of $P_{\mu, v}$ with $T^*(G/G_v)$ by Theorem 4.33 of Abraham and Marsden [1978].

Remark. In general, the symplectic form on $T^*(G/G_v)$ is the canonical one plus the "magnetic" field $d\alpha_\mu$. For the examples in this paper $d\alpha_\mu = 0$. However $\alpha_\mu \neq 0$ and it is necessary to use it to form the amended potential. If this is added to the standard potential for the heavy top, we recover the effective potential. As we shall see below, this agrees with that in standard texts (Goldstein [1980], formula 5-60, p. 215) and is a special case of Theorem 4.5.6 of Abraham and Marsden [1978].

Now consider the semi-direct product $G \times \mathfrak{o}_v^*$ of G with the additive group \mathfrak{o}_v^* with G acting on \mathfrak{o}_v^* by the adjoint action. For $v \in \mathfrak{o}_v^*$, let

$$f^v : G \times \mathfrak{g}^* \rightarrow \text{IR}$$

$$f^v(g, \xi) = \langle v, \text{Ad}_{g^{-1}} \xi \rangle \quad (2.10)$$

i.e. $f(g, \cdot) = \text{Ad}_{g^{-1}}^* v$. Write $f_\xi^v(g) = f^v(g, \xi)$. From the identity

$$f^v(g_1^{-1}g, \xi) = f^v(g, \text{Ad}_{g_1} \xi) \quad (2.11)$$

we see that $G \times \mathfrak{g}^*$ acts on T^*G by the right action

$$(g, \xi) \cdot \alpha_h = t_{\frac{df^v}{d\xi}} \cdot (\text{TR}_g)^* \cdot \alpha_h \quad (2.12)$$

where $t_{\frac{df^v}{d\xi}}$ is fiber translation in T^*G by the differential of f_ξ^v . The action (2.12) is symplectic and has an Ad^* -equivariant momentum map

$$\Lambda^v : T^*G \rightarrow \mathfrak{g}^* \times \mathfrak{g}^*$$

given by

$$\Lambda^v(\alpha_g) = ((T_e L_{g^{-1}})^* \alpha_g, \text{Ad}_{g^{-1}}^* v) \quad (2.13)$$

We shall write $m = (T_e L_{g^{-1}})^* \alpha_g$ and $v = \text{Ad}_{g^{-1}}^* v$. (See Abraham and Marsden [1978], Ex. 4.2c and Guillemin and Sternberg [1980]). We note that the adjoint action of $G \times \mathfrak{g}^*$ on $\mathfrak{g}^* \times \mathfrak{g}^*$ is

$$\text{Ad}_{(g, \xi)}(n, \zeta) = (\text{Ad}_g n, \text{Ad}_g \zeta + [\xi, \text{Ad}_g n]) \quad (2.14)$$

and that the Lie bracket on $\mathfrak{g}^* \times \mathfrak{g}^*$ is

$$[(\xi_1, n_1), (\xi_2, n_2)] = ([\xi_1, \xi_2], [\xi_1, n_2] + [n_1, \xi_2]) \quad (2.15)$$

One can check directly from (2.14) and (2.12) that Λ^v is equivariant.

(The identity

$$\text{Ad}_g[\xi, \eta] = [\text{Ad}_g \xi, \eta]$$

is useful in verifying this).

Observe that Λ^ν is invariant under the left action of G_ν ; i.e. for $h \in G_\nu$

$$\Lambda^\nu(TL_{h^{-1}}^* \alpha_g) = \Lambda^\nu(\alpha_g) \quad (2.16)$$

Thus, Λ^ν induces a map

$$\bar{\Lambda}^\nu : P_{\mu, \nu} \rightarrow \mathcal{O}_\mu^* \times \mathcal{O}_\nu^*$$

One readily sees that $\bar{\Lambda}^\nu$ is a diffeomorphism of $P_{\mu, \nu}$ onto the orbit $O_{\mu, \nu}$.

2.1 Theorem $\bar{\Lambda}^\nu : P_{\mu, \nu} \rightarrow O_{\mu, \nu}$ is a symplectic diffeomorphism.

Remarks. 1. This result is due to Ratiu [1981]. The proof we give is simpler, being motivated by Guillemin and Sternberg [1980].

2. For the Lagrange top, 2.1 can be proved by a direct, but messy, calculation which we outline in the next section.

Proof of Theorem 2.1. Since the symplectic form on $P_{\mu, \nu}$ is induced from the canonical symplectic structure on T^*G and that on $O_{\mu, \nu}$ is determined by the KKS structure, it suffices to show that Λ^ν commutes with Poisson brackets. This, however, is a general fact about Ad^* -equivariant momentum maps and collective Hamiltonians proved in the next two lemmas.

2.2 Lemma. Let $J : P \rightarrow \mathcal{O}^*$ be an Ad^* equivariant momentum map for the right action of a Lie group K on the symplectic manifold P . Let $F : \mathcal{O}^* \rightarrow \mathbb{R}$.

Then the Hamiltonian vector field $X_{F \circ J}$ for the (collective) Hamiltonian $F \circ J : P \rightarrow \mathbb{R}$ is given at $x \in P$ by

$$X_{F \circ J}(x) = \left[\frac{\delta F}{\delta \mu} \right]_P(x) \quad (2.17)$$

where $\frac{\delta F}{\delta \mu} \in \mathcal{L}$ is evaluated at $\mu = J(x)$ and ξ_P stands for the vector field on P generated by the Lie algebra element ξ .

Proof. By the chain rule and definition of the functional derivative,

$$\begin{aligned} d(F \circ J) \cdot v_x &= dF(\mu) \cdot dJ(x) \cdot v_x \\ &= \langle dJ(x) \cdot v_x, \frac{\delta F}{\delta \mu} \rangle \end{aligned}$$

where $v_x \in T_x P$ and $\mu = J(x)$ as above. By definition of a momentum map,

$$\langle dJ(x) \cdot v_x, \frac{\delta F}{\delta \mu} \rangle = \omega_x \left[\left[\frac{\delta F}{\delta \mu} \right]_P, v_x \right]$$

where ω is the symplectic form. Thus

$$d(F \circ J) \cdot v_x = \omega_x \left[\left[\frac{\delta F}{\delta \mu} \right]_P, v_x \right]$$

which means $X_{F \circ J} = \left[\frac{\delta F}{\delta \mu} \right]_P$. \blacksquare

2.3 Lemma. Under the assumptions of the preceding lemma, if $F, G : \mathcal{L}^* \rightarrow \mathbb{R}$, then

$$\{F \circ J, G \circ J\} = \{(F, G)\} \circ J \quad (2.18)$$

Proof. $\{F \circ J, G \circ J\} = \omega_x \left[\left[\frac{\delta F}{\delta \mu} \right]_P, \left[\frac{\delta G}{\delta \mu} \right]_P \right]$ $\quad (2.19)$

by definition of Poisson brackets on P and the preceding lemma. On the other hand, the KKS bracket is

$$\{\{F,G\}\}(\mu) = -\langle J(x), [\frac{\delta F}{\delta \mu}, \frac{\delta G}{\delta \mu}] \rangle \quad (2.20)$$

On the other hand, Ad^* equivariance gives (Abraham and Marsden [1978, Corollary 4.2.9])

$$\langle J(x), [\frac{\delta F}{\delta \mu}, \frac{\delta G}{\delta \mu}] \rangle = -\{ \langle J(x), \frac{\delta F}{\delta \mu} \rangle, \langle J(x), \frac{\delta G}{\delta \mu} \rangle \} \quad (2.21)$$

in which μ is fixed and x is variable. There is a minus sign here since J is generated by a right action. Holding $\frac{\delta F}{\delta \mu}$ fixed,

$$d\langle J(x), \frac{\delta F}{\delta \mu} \rangle \cdot v_x = v_x = \langle dJ(x) \cdot v_x, \frac{\delta F}{\delta \mu} \rangle$$

so $\langle J(x), \frac{\delta F}{\delta \mu} \rangle$ generates the same Hamiltonian vector field as $F \circ J$, as in 2.2. Thus, substitution of (2.21) in (2.20) yields (2.19). ■

Remarks. As we shall see, the heavy rigid top Hamiltonian is a collective Hamiltonian for J^V . This is compatible with the reduction picture: a Hamiltonian system on the reduced space $P_{\mu, v}$ can be written in the form

$$F = \{\{F, H\}\} \quad (2.22)$$

where H is written in terms of the KKS variables $(m, v) \in \mathcal{O}_f^* \times \mathcal{O}_f^*$.

§3. The rigid body

This section is divided into three parts. The first part explains how to write the equations of a free rigid body free to rotate about its center of mass in KKS form (2.5). The second part recalls the Euler angle formulation of the heavy top and the third part puts it into KKS form giving the explicit relationships with Euler angles.[†]

(a) The free rigid body

The free rigid body is a left invariant Hamiltonian system on $T^*SO(3)$, where $SO(3)$ is the group of proper orthogonal linear transformations of \mathbb{R}^3 to itself. By general facts about reduction we know that the equations of motion must be in the form (2.5): this is true of any left invariant Hamiltonian system on a Lie group G . For the free rigid body we can bypass the Euler angle description (the relevant formulas are given in Table 1 below).

The Lie algebra $so(3)$ of $SO(3)$ consists of the set of 3×3 skew symmetric matrices. We identify $so(3)$ with \mathbb{R}^3 by identifying

$$v = (p, q, r) \in \mathbb{R}^3 \text{ with } \hat{v} = \begin{bmatrix} 0 & -r & q \\ r & 0 & -p \\ -q & p & 0 \end{bmatrix} \in so(3) \quad (3.1)$$

The Lie bracket corresponds to the cross product in the sense that

$$[\hat{v}, \hat{w}] = (\hat{v} \times \hat{w})^\wedge. \quad (3.2)$$

[†]The third part may be omitted if desired since we give two proofs of our main result for the heavy top, one using only Euler angles and the other using the KKS structure

We denote elements of $so(3)^*$ by m ; these will also be identified with elements of \mathbb{R}^3 . Elements $m \in so(3)^*$ represent the body angular momentum of the rigid body and are related to the angular velocity ω by

$$m_i = I_i \omega_i, \quad i = 1, 2, 3 \quad (3.3)$$

where I_i are the moments of inertia. (The moment of inertia tensor has, as usual, been diagonalized). The standard Euler equations (Goldstein [1980], p. 205) written in terms of m are

$$\left. \begin{aligned} \dot{m}_1 &= \frac{I_2 - I_3}{I_2 I_3} m_2 m_3 \\ \dot{m}_2 &= \frac{I_3 - I_1}{I_1 I_3} m_1 m_3 \\ \dot{m}_3 &= \frac{I_1 - I_2}{I_1 I_2} m_1 m_2 \end{aligned} \right\} \quad (3.4)$$

Taking the Hamiltonian to be

$$H(m) = \frac{1}{2} \sum_{j=1}^3 \frac{m_j^2}{I_j} \quad (3.5)$$

we see by a simple calculation that (3.4) are equivalent to

$$\dot{F} = \{ \{ F, H \} \} \quad (3.6)$$

where $\{ \cdot, \cdot \}$ is the KKS bracket. In the present case this bracket becomes

$$\{ \{ F, G \} \} (m) = -m \cdot (\nabla F \times \nabla G) \quad (3.7)$$

for F, G functions of m .

The fact that the equations (3.6) must preserve coadjoint orbits amounts in this case to the fact that

$$\ell^2 = m_1^2 + m_2^2 + m_3^2 \quad (3.8)$$

is an (obvious) constant of the motion for (3.4). In terms of coadjoint orbits, equations (3.4) are Hamiltonian on each sphere in m -space with Hamiltonian function (3.5).[†]

The flow lines are given by intersecting the ellipsoids $H = \text{constant}$ with the spheres. For distinct moments of inertia the flow on the sphere has saddle points at $(0, \pm\ell, 0)$ and centers at $(\pm\ell, 0, 0)$, $(0, 0, \pm\ell)$. The saddles are connected by four heteroclinic orbits, as indicated in Figure 1.

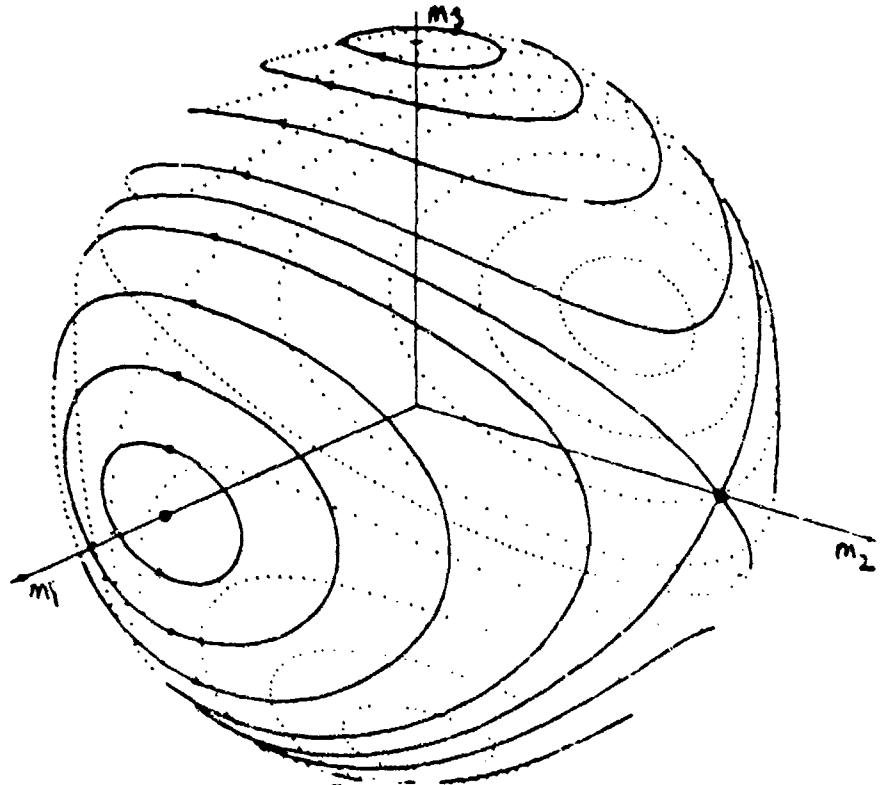


Figure 1. The spherical phase space of the rigid body for fixed total angular momentum $\ell = \sqrt{m_1^2 + m_2^2 + m_3^2}$; $I_1 > I_2 > I_3$

[†]The coadjoint orbits are spheres because the coadjoint action of $SO(3)$ on m space is just by rotations. The KKS symplectic structure is proportional to the area element by a simple calculation.

The orbits are, of course, explicitly known in terms of elliptic functions. The orbits of the most interest to us are the heteroclinic orbits which are given as follows. The four orbits lie in the invariant planes

$$m_3 = \pm \sqrt{\frac{a_3}{a_1}} m_1$$

where

$$a_1 = \frac{I_2 - I_3}{I_2 I_3} > 0, \quad a_2 = \frac{I_3 - I_1}{I_1 I_3} < 0 \quad \text{and} \quad a_3 = \frac{I_1 - I_2}{I_1 I_2} > 0.$$

3.1 Theorem. The heteroclinic orbits for the free rigid body are given by

$$\begin{aligned} m_1^+(t) &= \pm \sqrt{\frac{a_1}{-a_2}} \operatorname{sech}(-\sqrt{a_1 a_3} \ell t), \\ m_2^+(t) &= \pm \sqrt{\frac{a_1}{-a_2}} \tanh(-\sqrt{a_1 a_3} \ell t), \\ m_3^+(t) &= \pm \sqrt{\frac{a_3}{-a_2}} \operatorname{sech}(-\sqrt{a_1 a_3} \ell t), \end{aligned} \quad (3.9)$$

for $m_3 = +\sqrt{\frac{a_3}{a_1}} m_1$ and by

$$m_1^-(t) = m_1^+(-t), \quad m_2^-(t) = m_2^+(-t), \quad m_3^-(t) = -m_3^+(-t) \quad (3.10)$$

for $m_3 = -\sqrt{\frac{a_3}{a_1}} m_1$.

This may be checked by direct computation or by consulting one of the classical texts.

(b) The Heavy Top: Euler Angle Description

We now recall the traditional Euler angle description of the heavy top and shall locate homoclinic orbits for the symmetric (Lagrange) top. The Euler angle description is more familiar but in some respects the KKS description is simpler. For this reason we shall present both.

Given a rotation $A \in SO(3)$ we let the corresponding Euler angles be denoted (ϕ, ψ, θ) using the conventions of Goldstein [1980]; see Figure 2(a). The corresponding conjugate momenta are denoted p_ϕ, p_ψ, p_θ so that $(\phi, \psi, \theta, p_\phi, p_\psi, p_\theta)$ coordinatize $T^*SO(3)$.

We let \mathbf{m} denote the angular momentum in the body and let $\mathbf{v} = A^{-1}\mathbf{k}$ where \mathbf{k} is the unit vector along the spatial z-axis. We assume the center of mass is at $(0, 0, l)$ when A is the identity. The vectors (\mathbf{m}, \mathbf{v}) are expressed in the body coordinate system; see Figure 2(b).

The Hamiltonian is

$$H(\mathbf{m}, \mathbf{v}) = \frac{1}{2} \sum_{j=1}^3 \frac{\mathbf{m}_j^2}{I_j} + Mg\mathbf{l} \cdot \mathbf{v}_3 \quad (3.11)$$

where M is the total mass. When written in terms of Euler angles, this becomes

$$H = \frac{1}{2} \left\{ \frac{[(p_\phi - p_\psi \cos \theta) \sin \psi + p_\theta \sin \theta \cos \psi]^2}{I_1 \sin^2 \theta} + \frac{[(p_\theta - p_\psi \cos \theta) \cos \psi - p_\phi \sin \theta \sin \psi]^2}{I_2 \sin^2 \theta} + \frac{p_\psi^2}{I_3} \right\} + Mg\mathbf{l} \cdot \mathbf{v}_3 \quad (3.12)$$

In Table 1 below we summarize the relationships between $\mathbf{m}, \mathbf{v}, \phi, \psi, \theta, p_\phi, p_\psi, p_\theta, \dot{\phi}, \dot{\psi}$ and $\dot{\theta}$ for convenient reference, but in this subsection we shall take (3.12), or equivalently the Lagrangian

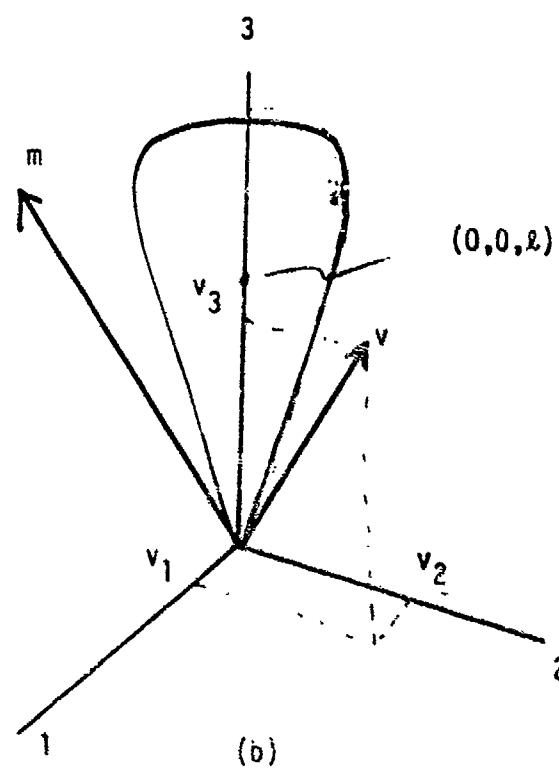
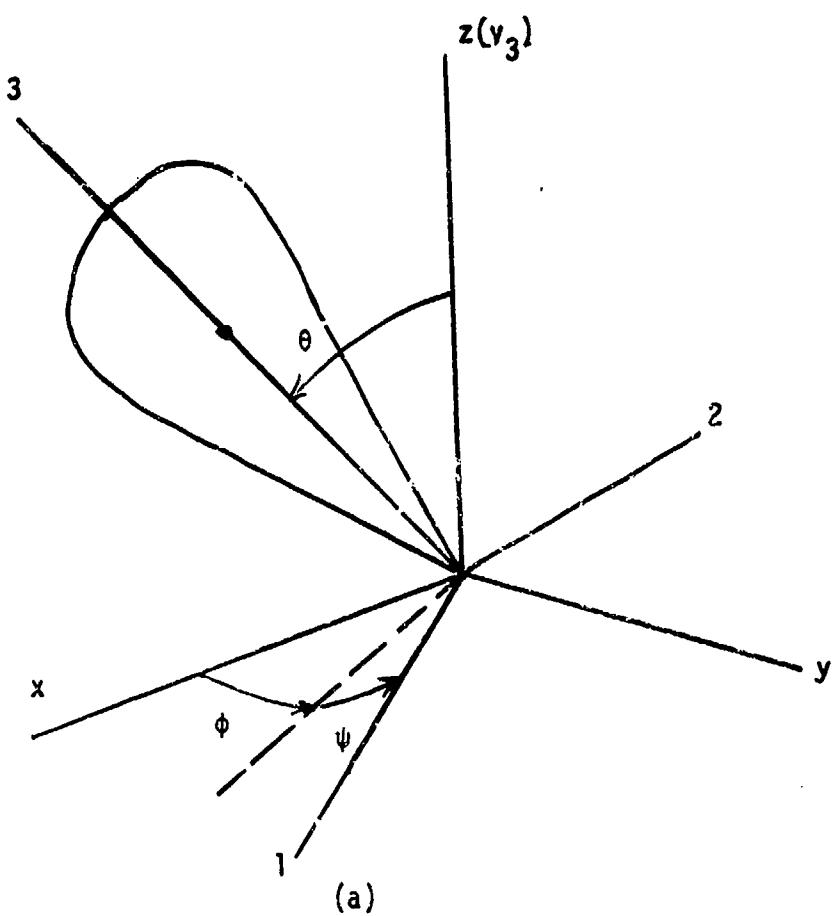


Figure 2. The heavy rigid body, illustrating space (x,y,z) and body $(1,2,3)$ coordinates, and the Euler angles (ϕ, ψ, θ) .

$$\begin{aligned}
 L = & \frac{I_1}{2} (\dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi)^2 + \frac{I_2}{2} (\dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi)^2 \\
 & + \frac{I_3}{2} (\dot{\phi} \cos \theta + \dot{\psi})^2 - Mg\ell \cos \theta
 \end{aligned} \tag{3.13}$$

as our starting point, and work exclusively in terms of Euler angles.

The Hamiltonian (3.12) is invariant under rotations about the z-axis; i.e. ϕ is a cyclic variable, so p_ϕ is a constant of the motion. In other terms, the momentum map for this S^1 action is $J(\phi, \psi, \theta, p_\phi, p_\psi, p_\theta) = p_\phi$. The reduced space $J^{-1}(p_\phi)/S^1$ is parametrized by $(\psi, \theta, p_\psi, p_\theta)$. In fact this reduced space is identifiable with T^*S^2 , the cotangent bundle of the two sphere on which $(\psi, \theta, p_\psi, p_\theta)$ are canonical coordinates. In fact, as we shall see below, (ψ, θ) are a system of spherical coordinates for this sphere. The equations of motion for ψ, θ are thus just Hamilton's equations for (3.12) with p_ϕ held constant.

Remarks 1. The S^1 reduction here is in accord with the general fact that reducing T^*Q by S^1 gives $T^*(Q/S^1)$; here $Q = SO(3)$ and $SO(3)/S^1 = S^2$.

2. The two sphere obtained here is not to be confused with the sphere for the free rigid body shown in Figure 1.

For the symmetric top, $I_1 = I_2$, ψ is also a cyclic variable and p_ψ is constant as well. In fact these two S^1 symmetries commute, so we have a torus, T^2 , symmetry, which makes the system completely integrable. The reduced system has one degree of freedom in (θ, p_θ) . The reduced space is no longer a manifold, but has a boundary at $\theta = 0, \pi$; i.e. it is $T^*[0, \pi]$.

These singularities in the reduced space correspond to the fact that the level set $p_\psi = \text{constant}$ is singular at $\theta = 0, \pi^+$.

With $I_1 = I_2$, (3.12) becomes

$$\begin{aligned} H &= \frac{1}{2I_1 \sin^2 \theta} \{ (p_\phi - p_\psi \cos \theta)^2 + p_\theta^2 \sin^2 \theta \} + \frac{p_\psi^2}{I_3} + Mg\ell \cos \theta \\ &= \frac{p_\theta^2}{2I_1} + \left(Mg\ell \cos \theta + \frac{(p_\phi - p_\psi \cos \theta)^2}{2I_1 \sin^2 \theta} + \frac{p_\psi^2}{I_3} \right) \end{aligned} \quad (3.14)$$

which shows the amended potential explicitly.

Remark. In these coordinates, the abstract formula (2.9) for the ϕ, ψ -reductions gives the closed one forms $p_\phi dp_\phi$ and $p_\psi dp_\psi$ respectively.

3.2 Theorem. If $0 < p < 2\sqrt{Mg\ell I_1}$ and if $p_\phi = p_\psi = I_1 b$, the reduced Lagrange top system in (θ, p_θ) space has a hyperbolic saddle points at $\theta = 0$, $p_\theta = 0$ and a homoclinic orbit connecting it to itself given by

$$\cos \theta = 1 - \gamma \operatorname{sech}^2 \left(\frac{\sqrt{B\gamma}}{2} t \right) \quad (3.15)$$

where $B = 2 Mg\ell / I_1$ and $\gamma = 2 - \frac{b^2}{B}$.

Remarks 1. A top with $0 < p_\psi < 2\sqrt{Mg\ell I_1}$ for which the vertically spinning state is unstable (a saddle) is called a slow top.

[†]This is a special case of a general fact about singularities in level sets of momentum maps; see Arms, Marsden and Moncrief [1981].

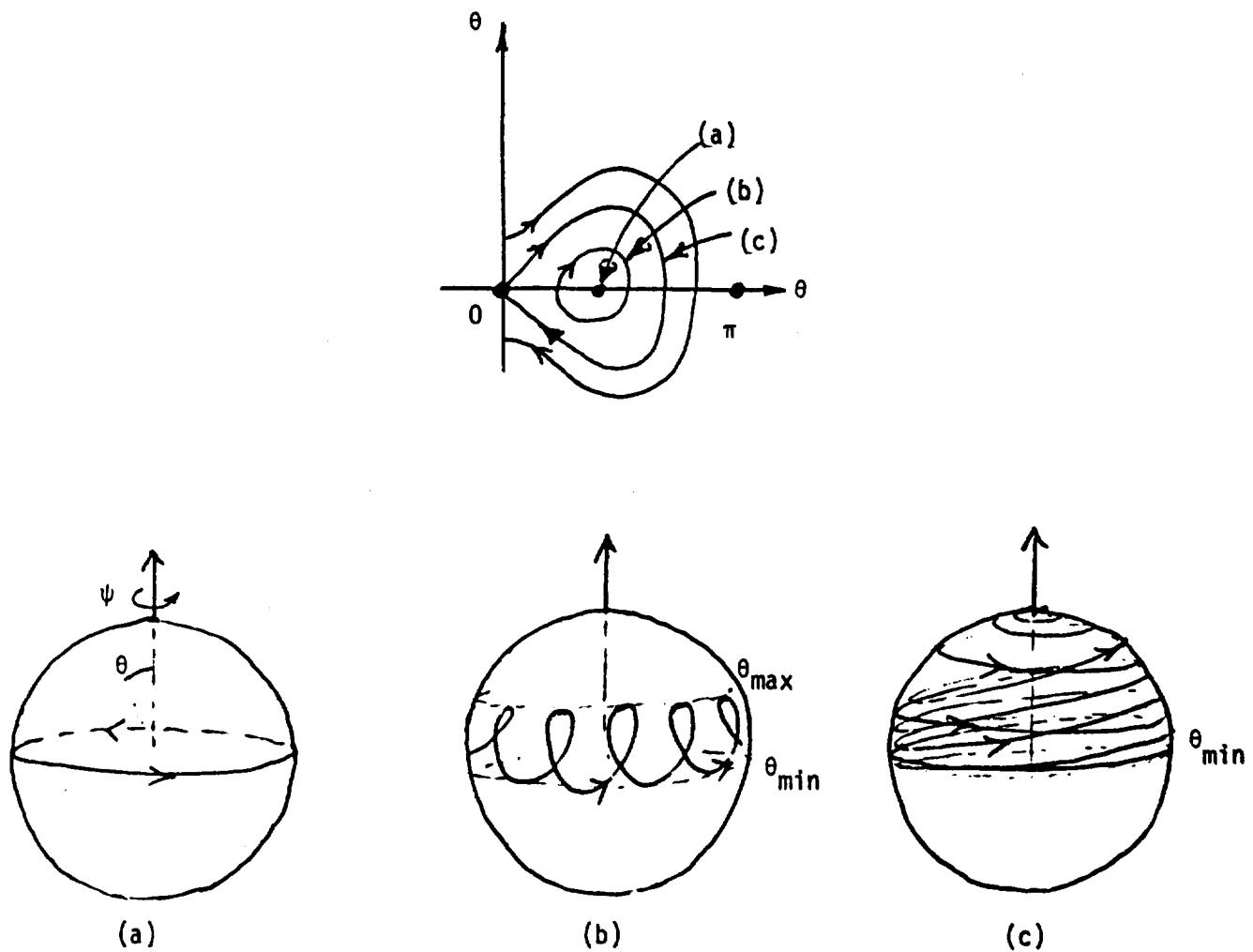


Figure 3. Motions of the top a) steady precession ($\theta = \text{constant}$)
b) precession and notation ($\theta_{\min} \leq \theta \leq \theta_{\max}$)
c) homoclinic orbit ($\theta_{\min} \leq \theta < 0$)

2. For $p_\phi = p_\psi$ note that the potential in (3.14) does not have a singularity at $\lambda = 0, \pi$, so these end points do not cause difficulties.

3. Note that $0 < \gamma < 2$.

The theorem is most easily checked by using the energy equation (3.14); the homoclinic orbit has energy $H = Mgl + P_\psi^2/2I_3$; cf. Goldstein [1980], pp. 215, 216 and Figure 3.

(c) The heavy top: KKS description

The abstract theory in the preceding section guarantees that the reduced space T^*S^2 for the heavy top is symplectically diffeomorphic to a coadjoint orbit in the semi-direct product $SO(3) \times \mathbb{R}^3$; i.e. in the Euclidean group E_3 . The Lie algebra is denoted e_3 .

The mapping that gives this diffeomorphism is just the map

$$\Lambda : (\phi, \psi, \theta, p_\phi, p_\psi, p_\theta) \mapsto (m, v) \quad (3.16)$$

where m is the angular momentum in the body and v is the orientation of gravity as viewed from the body. Table 1 below summarizes the explicit formulas relating these quantities. Table 2 summarizes the relationships between the "Euler angle" spaces and the coadjoint spaces.

The KKS bracket for functions of (m, v) is given by (2.2) and (2.15), which in this case becomes

$$\{ \{ F, G \} \} (m, v) = -m \cdot (\nabla_m F \times \nabla_m G) - v \cdot (\nabla_m F \times \nabla_v G + \nabla_v F \times \nabla_m G) \quad (3.17)$$

The assertions of the general theory can be checked by hand in this case. First of all, it is easily seen that the KKS bracket equation $\dot{F} = \{ \{ F, H \} \}$ with H given by (3.11) and the bracket by (3.17) yield the equations

Table 1

Formulas Relating Euler Variables and KKS Variables for the Heavy Top

$$m_1 = [(p_\phi - p_\psi \cos \theta) \sin \psi + p_\theta \sin \theta \cos \psi] / \sin \theta = I_1(\dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi)$$

$$m_2 = [(p_\phi - p_\psi \cos \theta) \cos \psi - p_\theta \sin \theta \sin \psi] / \sin \theta = I_2(\dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi)$$

$$m_3 = p_\psi = I_3(\dot{\phi} \cos \theta + \dot{\psi})$$

$$v_1 = \sin \theta \sin \psi$$

$$v_2 = \sin \theta \cos \psi$$

$$v_3 = \cos \theta$$

$$p_\phi = m \cdot v = I_1(\dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi) \sin \theta \sin \psi + I_2(\dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi) \sin \theta \cos \psi$$

$$p_\psi = m_3 = I_3(\dot{\phi} \cos \theta + \dot{\psi})$$

$$p_\theta = (v_2 m_1 - v_1 m_2) / \sqrt{1 - v_2^2} = I_1(\dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi) \cos \psi - I_2(\dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi) \sin \psi$$

$$\dot{\phi} = \frac{1}{I_1} \frac{m_1 v_1}{\sqrt{1 - v_3^2}} + \frac{1}{I_2} \frac{m_2 v_2}{\sqrt{1 - v_3^2}}$$

$$\dot{\psi} = \frac{m_3}{I_3} - \frac{m_3 m_1 v_1}{I_1 \sqrt{1 - v_3^2}} - \frac{m_3 m_2 v_2}{I_2 \sqrt{1 - v_3^2}}$$

$$\dot{\theta} = \frac{m_1 v_2}{I_1 \sqrt{v_1^2 + v_2^2}} - \frac{m_2 v_1}{I_2 \sqrt{v_1^2 + v_2^2}}$$

Table 2
The Relationship Amongst the Spaces and Variables

Euler Angles		KKS	
Variables	Space	Space	Variables
$(\phi, \psi, \theta, p_\phi, p_\psi, p_\theta)$	$T^*SO(3)$ ↓ reduce by $S^1 : \phi$ cyclic	e_3^* ↓ restriction	(m, v)
$(\psi, \theta, p_\psi, p_\theta)$	T^*S^2 ↓ If $I_1 = I_2$, Lagrange reduction	coadjoint orbit	$\ v\ = 1, m \cdot v = p_\phi$
(θ, p_θ)	$T^*[0, \pi]$	coadjoint orbit/ S^1	$v_2 m_1 - v_1 m_2, v_3$

$$\left. \begin{array}{l} \dot{m}_1 = a_1 m_2 m_3 - M g \ell v_2 \\ \dot{m}_2 = a_2 m_1 m_3 + M g \ell v_1 \\ \dot{m}_3 = a_3 m_1 m_2 \end{array} \right\} \quad (3.18)$$

$$\left. \begin{array}{l} \dot{v}_1 = \frac{m_3 v_1}{I_3} - \frac{m_2 v_3}{I_2} \\ \dot{v}_2 = \frac{m_1 v_3}{I_1} - \frac{m_3 v_1}{I_3} \\ \dot{v}_3 = \frac{m_2 v_1}{I_2} - \frac{m_1 v_2}{I_1} \end{array} \right\} \quad (3.19)$$

A slightly tedious, though straightforward computation shows that these are equivalent to the Hamiltonian equations for (3.12). In fact, if F, G are functions of (m, v) and (m, v) are then written out in terms of the Euler variables then an explicit (but tedious) computation shows that

$$\{\{F, G\}\}(m, v) = \{F, G\}_{\theta, \psi} \quad (3.20)$$

where $\{ , \}_{\theta, \psi}$ denotes the bracket computed for the θ, ψ variables (i.e. holding p_ϕ constant). This is the content of Theorem 2.1 for this special case.

The equations (3.18) and (3.19) have $\|v\|$ and $m \cdot v$ as constants of the motion. This just reflects (a) the conservation law $p_\phi = \text{constant}$ and (b) the preservation of the coadjoint orbit by the KKS equations. The conditions $\|v\| = 1$ and $m \cdot v = p_\phi = \text{constant}$ also give the explicit identification of the coadjoint orbit with $T^* S^2$. Indeed, $\|v\| = 1$ describes the unit sphere S^2 and $m \cdot v = p_\phi$ specifies m as a linear functional on the unit normal to S^2 , leaving m restricted to $T_v S^2$ free. Thus m determines,

by restriction, an element of $T_v^*S^2$. Finally observe that the equations for v in Table 1 show that θ, ψ are spherical-type coordinates on S^2 (ψ has been rotated by $\pi/2$ from standard conventions on spherical coordinates).

Finally we discuss the Lagrange top in the KKS picture. For $I_1 = I_2$ the invariance is rotations about the 3 axis. This S^1 action corresponds to the S^1 action of rotation through ψ in the Euler angle picture, as is easily seen. Also, the momentum map can be directly checked in the KKS picture to be just m_3 .

The following is a general property of reduction, but it may also be checked explicitly in this case.

3.3 Lemma. If F and G are functions of (m, v) which are rotationally invariant, then

$$\{\{F, G\}\} = \{\{F, G\}\}_{m_3} = \text{const} \quad (3.21)$$

i.e. the full KKS bracket is the same as if it were computed holding m_3 constant.

With $\|v\| = 1$, $m \cdot v = p_\phi$, any rotationally invariant function of (m, v) can in fact be expressed as a function of $v_2 m_1 - v_1 m_2$ and v_3 . These variables are a convenient representation of the coadjoint space reduced by the S^1 action. Brackets of functions of these variables may be computed by 3.3. Again, one can check by hand that for such functions

$$\{\{F, G\}\} = \{F, G\}_\theta \quad (3.22)$$

where

$$\{F, G\}_\theta = \frac{\partial F}{\partial \theta} \frac{\partial G}{\partial p_\theta} - \frac{\partial G}{\partial \theta} \frac{\partial F}{\partial p_\theta}$$

i.e. the canonical bracket holding p_ϕ, p_ψ constant.

The homoclinic orbit now is described very simply as follows:

3.4 Theorem. For $0 < m_3 < 2\sqrt{Mg\ell} I_1$ and $m \cdot v = m_3$, the reduced system for the heavy top has a hyperbolic saddle point at

$$v_3 = 0, \quad v_2 m_1 - v_1 m_2 = 0$$

and a homoclinic orbit connecting it to itself given by

$$\left. \begin{aligned} v_3 &= 1 - \gamma \operatorname{sech}^2 \left(\frac{\sqrt{\beta\gamma}}{2} t \right) \\ v_2 m_1 - v_1 m_2 &= m_3 (1 - v_3) \sqrt{1 - v_3^2} \quad [= I_1 \dot{\phi} \sin^2 \theta] \end{aligned} \right\} \quad (3.23)$$

and

where β, γ are as in Theorem 3.2 with $p_\psi = m_3$.

This can be checked directly (using the energy equation and the evolution equations). Of course it corresponds to (3.15).

§4. Transversal intersections of invariant manifolds for KKS variables coupled to action angle variables

We now develop an extension of the Melnikov-Arnold theory (Melnikov [1963], Arnold [1964], Holmes [1980]) which applies to systems described by a set of m KKS variables $\mu = (\mu_1, \dots, \mu_m) \in \mathcal{O}_f^*$ and a set of n action angle variables $(\theta_1, I_1), \dots, (\theta_n, I_n)$. This theory will then be applied to the free rigid body with attachments in the next section. For the heavy top, a more sophisticated version is needed for systems whose variables do not decompose so cleanly. This is the subject of Section 6.

The action angle variables can be those associated with motion near an elliptic fixed point in a one degree of freedom system. For example, our methods apply to the system obtained by coupling two rigid bodies if we examine the motion near a homoclinic orbit in one (such as an orbit connecting $(0, \ell, 0)$ to $(0, -\ell, 0)$ in Figure 1) and a small periodic orbit in the other (such as an orbit near $(\ell, 0, 0)$ in Figure 1). The coupled system then will have horseshoes under the conditions of Theorem 4.3 below. (We believe that this applies, in particular, to the five mode truncation of the Euler equations for an ideal fluid on a two torus, which consists of two sets of overlapping and coupled rigid body equations; note that for $\mathcal{O}_f = \text{so}(3)$ and $n = 1$, we have five variables in the set (μ, θ, I) .) For simplicity, however, we have chosen a rigid body with attachments to work out in detail since the action angle variables are more explicit and it is easy to add on additional ones.

We assume that our Hamiltonian takes the form

$$\begin{aligned}
 H^\epsilon(\mu, \theta, I) &= F(\mu) + \sum_{j=1}^n G_j(I_j) + \epsilon H^1(\mu, \theta, I) \\
 &= H^0(\mu, I_1, \dots, I_n) + \epsilon H^1(\mu, \theta, I) + O(\epsilon^2)
 \end{aligned} \tag{4.1}$$

where $\mu = (\mu_1, \dots, \mu_m) \in \mathfrak{o}_j^*$, the dual space of a Lie algebra \mathfrak{o}_j and $\theta = (\theta_1, \dots, \theta_n)$, $I = (I_1, \dots, I_n)$ with θ_j a 2π -periodic variable. We assume that the KKS system associated with F has a homoclinic (or heteroclinic) orbit $\bar{\mu}(t) \in \mathfrak{o}_j^*$. The oscillator frequencies

$$\Omega_j(I_j) = \frac{\partial G_j}{\partial I_j} \quad (4.2)$$

are assumed to be positive.

As in Holmes and Marsden [1981b,c] we can solve the equation

$$H^\varepsilon(\mu, \theta, I) = h \quad (4.3)$$

for I_n in the form

$$I_n = L^\varepsilon(\mu, \theta_1, \dots, \theta_n, I_1, \dots, I_{n-1}, h) = L^0(\mu, I_1, \dots, I_{n-1}, h) + \varepsilon L^1(\mu, \theta_1, \dots, \theta_n, I_1, \dots, I_{n-1}, h) + O(\varepsilon^2) \quad (4.4)$$

where

$$L^0(\mu, I_1, \dots, I_{n-1}, h) = G_n^{-1}(h - F(\mu) - \sum_{j=1}^{n-1} G_j(I_j)) \quad (4.5)$$

and

$$L^1(\mu, \theta_1, \dots, \theta_n, I_1, \dots, I_{n-1}, h) = -\frac{H^1(\mu, \theta_1, \dots, \theta_n, I_1, \dots, I_{n-1}, L^0(\mu, I_1, \dots, I_{n-1}, h))}{\Omega_n(L^0(\mu, I_1, \dots, I_{n-1}, h))} \quad (4.6)$$

In addition, we eliminate t in favor of the new 'time' θ_n and write for $d/d\theta_n$.

4.1 Proposition. Hamilton's equation for (4.1), namely

$$\left. \begin{aligned} \dot{\mu}_i &= \{(\mu_i, H^\epsilon)\}, \quad i = 1, \dots, \\ \dot{\theta}_j &= \frac{\partial H^\epsilon}{\partial I_j}, \quad \dot{I}_j = -\frac{\partial H^\epsilon}{\partial \theta_j}, \quad j = 1, \dots, n-1 \end{aligned} \right\} \quad (4.7)$$

(where $\{ \cdot, \cdot \}$ denotes the KKS bracket in the μ -variables) become

$$\left. \begin{aligned} \dot{\mu}_i &= -\{(\mu_i, L^\epsilon)\} \\ \dot{\theta}_j &= -\frac{\partial L^\epsilon}{\partial I_j}, \quad \dot{I}_j = \frac{\partial L^\epsilon}{\partial \theta_j} \end{aligned} \right\} \quad (4.8)$$

Proof. Implicit differentiation of (4.3) gives

$$\frac{\delta H^\epsilon}{\delta \mu} + \frac{\delta H^\epsilon}{\delta I_n} \frac{\delta L^\epsilon}{\delta \mu} = 0$$

It follows that for a function $K(\mu)$,

$$\begin{aligned} \{(\{K, H^\epsilon\})(\mu)\} &= -\langle \mu, \left[\frac{\delta K}{\delta \mu}, \frac{\delta H^\epsilon}{\delta \mu} \right] \rangle \\ &= \langle \mu, \left[\frac{\delta K}{\delta \mu}, \frac{\delta H^\epsilon}{\delta I_n} \frac{\delta L^\epsilon}{\delta \mu} \right] \rangle \end{aligned}$$

and so $\{(\{K, H^\epsilon\})\} / \frac{\partial H^\epsilon}{\partial I_n} = -\{(\{K, L^\epsilon\})\}$. However then, $K' = \frac{\dot{K}}{\dot{\theta}_n} = \frac{\dot{K}}{\frac{\partial H^\epsilon}{\partial I_n}} = -\{(\{K, L^\epsilon\})\}$.

The second set of equations in (4.8) follows in the same way. ■

Using $L^\epsilon = L^0 + \epsilon L^1 + O(\epsilon^2)$ we see that (4.8) has the form of a periodically perturbed system. Next we relate the reduced and non-reduced brackets.

4.2 Proposition. We have

$$\{(L^0, L^1)\} = \frac{1}{(\Omega_n)^2} \{(F, H^1)\} \quad (4.9)$$

Proof. From (4.5),

$$\frac{\delta L^0}{\delta \mu} = - \frac{1}{\Omega_n} \frac{\delta F}{\delta \mu}$$

and

$$\frac{\delta L^1}{\delta \mu} = - \frac{1}{\Omega_n} \left(\frac{\delta H^1}{\delta \mu} + \frac{\partial H^1}{\partial T_n} \frac{\delta L^0}{\delta \mu} \right) + \frac{1}{(\Omega_n)^2} H^1 \frac{\partial \Omega_n}{\partial T_n} \frac{\delta L^0}{\delta \mu}$$

However, since $[\frac{\delta L^0}{\delta \mu}, \frac{\delta L^1}{\delta \mu}] = 0$, we get

$$\begin{aligned} \{(L^0, L^1)\} &= -\langle \mu, [\frac{\delta L^0}{\delta \mu}, \frac{\delta L^1}{\delta \mu}] \rangle \\ &= -\langle \mu, \frac{1}{\Omega_n^2} [\frac{\delta F}{\delta \mu}, \frac{\delta H^1}{\delta \mu}] \rangle \\ &= \frac{1}{\Omega_n^2} \{(F, H^1)\}. \blacksquare \end{aligned}$$

Let us now give a special case of the general result, suitable for two degree of freedom systems.

4.3 Theorem. Suppose $\bar{u}(t)$ is a homoclinic (or heteroclinic) orbit for F , which lies on a coadjoint orbit in \mathfrak{g}^* of dimension 2. Furthermore, suppose $n = 1$.

Let $\bar{h} = F(\bar{u})$ be the energy of the homoclinic orbit and let $h > \bar{h}$ and $\varepsilon^0 = G^{-1}(h - \bar{h})$ be constants. Let $\{(F, H^1)\}(t, \theta^0)$ denote the KKS bracket of $F(u)$ and $H^1(\mu, \Omega(\varepsilon^0)t + \theta^0, \varepsilon^0)$ evaluated at $\bar{u}(t)$. Let

$$M(\theta^0) = \frac{1}{\Omega(\ell^0)} \int_{-\infty}^{\infty} \{(F, H^1)\}(t, \theta^0) dt \quad (4.10)$$

and assume $M(\theta^0)$ has simple zeros. Then for $\epsilon > 0$ sufficiently small, the Hamiltonian system (4.1) contains transverse homoclinic orbits and hence Smale horseshoes on the energy surface $H^\epsilon = h$.

Proof. By reduction, it suffices to check that the Melnikov function for the reduced, forced system on the coadjoint orbit containing $\bar{\mu}$ has simple zeros. This involves only a generalization of the one degree of freedom Melnikov theory for forced oscillations to two dimensional symplectic manifolds. The standard proof (see Holmes [1980] and Greenspan and Holmes [1981]) carries over directly. The Melnikov function for the reduced system is

$$M(\theta^0) = \int_{-\infty}^{\infty} \{(\mathcal{L}^0, \mathcal{L}^1)\} d\theta \quad (4.11)$$

since the KKS bracket coincides with the Poisson bracket on coadjoint orbits. Using (4.9) and $\frac{d\theta}{dt} = \Omega$,

$$M(\theta^0) = \int_{-\infty}^{\infty} \frac{1}{\Omega} \{(F, H^1)\} dt$$

Finally, note that $\Omega = \Omega(\ell^0)$ is constant on the homoclinic orbit, so the theorem results. ■

To deal with the situation in which $n \geq 2$, we introduce the following conditions on the Hamiltonian (4.1).

(H1) F contains a homo- (or hetero) clinic orbit $\bar{\mu} \in \mathcal{M}^*$ with energy \bar{h} . The coadjoint orbit containing $\bar{\mu}$ is assumed to be two-dimensional. The saddle points for $\bar{\mu}$ are denoted μ_\pm (they could be coincident).

$$(H2) \quad \Omega_j(I_j) = G_j^1(I_j) > 0, \quad j = 1, \dots, n.$$

To explain the remaining conditions some discussion is needed.

For $\epsilon = 0$, note that the Hamiltonian system for L^0 has two $n-1$ parameter families of invariant $n-1$ dimensional tori $T_{\pm}(h_1, \dots, h_{n-1})$ given by

$$\left. \begin{aligned} \mu &= \mu_{\pm} \\ G_j(I_j) &= h_j = \text{constant (i.e. } I_j = \ell_j = G_j^{-1}(h_j)) \\ \theta_j &= \Omega_j(\ell_j)\theta_n + \theta_j^0 \pmod{2\pi}, \quad j = 1, \dots, n-1 \end{aligned} \right\} \quad (4.12)$$

(Correspondingly, the system for H^0 has two n parameter family of invariant tori $(T_{\pm}(h_1, \dots, h_n))$.) Henceforth we write the (phase) constants of integration θ_j^0 as θ_j^0 , $j = 1, \dots, n-1, n$.

The tori $T_{\pm}(h_1, \dots, h_{n-1})$ are connected by the n -dimensional homoclinic manifold defined by

$$\left. \begin{aligned} \mu &= \bar{\mu}(\theta_n - \theta_n^0) \\ G_j(I_j) &= h_j, \\ \theta_j &= \Omega_j(I_j)\theta_n + \theta_j^0, \quad j = 1, \dots, n-1, n \end{aligned} \right\} \quad (4.13)$$

where the phase constant θ_n^0 associated with the 'reduced' degree of freedom appears explicitly. This manifold consists of the coincident stable and unstable manifolds of the tori $T_{\pm}(h_1, \dots, h_{n-1})$; i.e.

$$W^s(T_{\pm}(h_1, \dots, h_{n-1})) = W^u(T_{\pm}(h_1, \dots, h_{n-1}))$$

given by (4.13).

For $\epsilon \neq 0$ the system (4.7) possesses a Poincaré map P_ϵ from (a piece of) $(\mu, \theta_1, \dots, \theta_{n-1}, I_1, \dots, I_{n-1})$ space to itself where θ_n goes through an increment of 2π , starting at some fixed value θ_n^0 , (which will be suppressed in the notation). Below, when we refer to transverse intersection of stable and unstable manifolds, we mean so for this Poincaré map.

(H3) Assume that the constants $G_j(I_j) = h_j$, $j = 1, \dots, n$ are chosen so that the unperturbed frequencies $\Omega_1(I_1), \dots, \Omega_n(I_n)$ satisfy the non-degeneracy conditions (i.e. $\Omega_j'(I_j) \neq 0$, $j = 1, \dots, n-1$) and the non-resonance conditions of the KAM theorem. (cf. Arnold (1978), Appendix 8).

This condition ensures that the tori $T_\pm(h_1, \dots, h_{n-1})$ perturb to invariant tori $T_{\epsilon, \pm}(h_1, \dots, h_{n-1})$ for P_ϵ for ϵ sufficiently small.

Let $h > \bar{h}$, $h = \bar{h} + \sum_{j=1}^n h_j$ where $h_j > 0$, and the unperturbed homoclinic manifold be filled with an n -parameter family of orbits given by $(\mu, \theta_1, \dots, \theta_n, I_1, \dots, I_n) = \bar{u}(t)$, $\Omega_1(I_1)t + \theta_1^0, \dots, \Omega_n(I_n)t + \theta_n^0, I_1, \dots, I_n$. Pick one such orbit and let $\{(F, H^1)\}$ denote the KKS Poisson bracket of $F(\mu)$ and $H^1(\mu, \theta_1, \dots, \theta_n, I_1, \dots, I_n)$ evaluated on this orbit. Similarly, let $\{I_k, H^1\} = -\partial H^1 / \partial \theta_k$, $k \in 1, \dots, n-1$ be evaluated on this orbit. Define the Melnikov Vector $M(\theta^0) = (M_1, \dots, M_{n-1}, M_n)$ by

$$M_k(\theta_1^0, \dots, \theta_n^0, h, h_1, h_2, \dots, h_{n-1}) = \int_{-\infty}^{\infty} \{I_k, H^1\} dt, \quad k = 1, \dots, n-1.$$

$$M_n(\theta_1^0, \dots, \theta_n^0, h, h_1, h_2, \dots, h_{n-1}) = \frac{1}{h_n} \int_{-\infty}^{\infty} \{(F, H^1)\} dt. \quad (4.14)$$

(We note that $h_n = h - \bar{h} - \sum_{j=1}^{n-1} h_j$; I_n and h_n do not explicitly enter the calculations, since I_n is eliminated by the reduction process; we also note that these integrals need not be absolutely convergent, but we do require conditional convergence.)

(H4) Assume that the multiply 2π periodic Melnikov vector $M : \mathbb{R}^n \rightarrow \mathbb{R}^n$ (which is independent of ϵ) has at least one transversal zero; i.e. there is a point $(\theta_1^0, \dots, \theta_n^0)$ for which

$$M(\theta_1^0, \dots, \theta_n^0) = 0$$

but

$$\det[DM(\theta_1^0, \dots, \theta_n^0)] \neq 0,$$

where DM is the $n \times n$ matrix of partial derivatives of M_1, \dots, M_n with respect to $\theta_1^0, \dots, \theta_n^0$, the initial phases of the orbit.

Here is the result for $n \geq 2$.

4.4 Theorem. If conditions (H1) - (H4) hold for the system (4.1), then, for ϵ sufficiently small, the perturbed stable and unstable manifolds $W^s(T_{\epsilon, \pm})$, and $W^u(T_{\epsilon, \mp})$ of the perturbed tori $T_{\epsilon, \pm}$ intersect transversely.

The theorem follows from the arguments of Holmes and Marsden [1981c] in the present context. We also refer the reader to that paper for a discussion of how this yields Arnold diffusion and for related references.

There is a similar result when the coadjoint orbit is higher dimensional i.e. of dimension $2M$, $M > 1$, but the system for F on \mathfrak{g}^* is completely

integrable, say with integrals

$$F = F_1, F_2, \dots, F_M.$$

where F_2, \dots, F_M are associated with action angle variables. Now a result similar to 4.4 holds if the Melnikov vector is enlarged by replacing the number

$$\frac{1}{\Omega_n} \int_{-\infty}^{\infty} \{\{F, H^1\}\} dt$$

by the vector

$$\frac{1}{\Omega_n} \int_{-\infty}^{\infty} \{\{F_i, H^1\}\} dt, i = 1, \dots, m$$

It would be of interest to apply such a generalization to the Toda lattice and related completely integrable systems.

35. An example: the rigid body with attachments

We start by considering the rigid body in the absence of gravity, with a single attachment which spins freely about an axis coincident with one of the body's principal axes, say 1. The angle about this axis is denoted θ . For simplicity we suppose that the attachment does not affect the position of the principal axes, so that the inertia tensor remains diagonal, but that the attachment is slightly asymmetrical in the sense that the perturbed moments of inertia about the 2 and 3 axes depend upon the angular position θ of the attachment relative to the body. Such an attachment could consist of a (heavy) disc and a light rod, as in Figure 4.

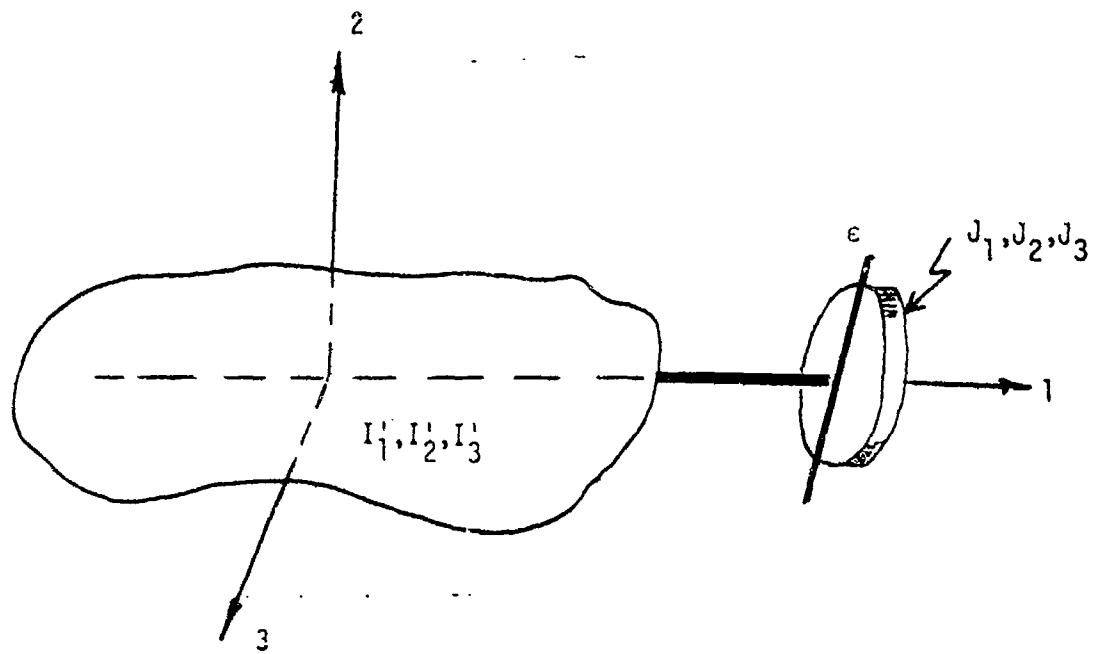


Figure 4

Letting the moments of inertia of the body with respect to the three body axes be $I_j^!$, $j = 1, 2, 3$, and of the attachment be $J_1 + \epsilon$, $J_2 + \epsilon \cos^2 \theta$, $J_3 + \epsilon \sin^2 \theta$. We obtain the Hamiltonian

$$\begin{aligned}
 H^\epsilon &= \frac{1}{2} \left\{ \frac{m_1^2}{I_1 + \epsilon} + \frac{m_2^2}{I_2 + \epsilon \cos^2 \theta} + \frac{m_3^2}{I_3 + \epsilon \sin^2 \theta} + \frac{I^2}{J_1 + \epsilon} \right\} \\
 &= \frac{1}{2} \sum_{j=1}^3 \frac{m_j^2}{I_j} + \frac{I^2}{2J_1} - \frac{\epsilon}{2} \left\{ \frac{m_1^2}{I_1^2} + \frac{m_2^2}{I_2^2} \cos^2 \theta + \frac{m_3^2}{I_3^2} \sin^2 \theta + \frac{I^2}{J_1^2} \right\} + O(\epsilon^2) \\
 &= F + G + \epsilon H^1 + O(\epsilon^2) \tag{5.1}
 \end{aligned}$$

where $I_j = I_j^1 + J_j$ and where (θ, I) are action angle variables for the attachment. Since the unperturbed system is a product flow on $(S^2) \times (IR \times S^1)$, the product of the coadjoint orbit of $SO(3)$ with the (I, θ) cylinder, we can use (3.9) to write the homoclinic orbits for an energy level $H^0 = h = \frac{1}{2} \left(\frac{\ell^2}{I_2} + \frac{k^2}{J_1} \right)$, where $\ell^2 = m_1^2 + m_2^2 + m_3^2$ and k is a constant, as

$$\begin{aligned}
 m_1 &= \pm \ell \sqrt{\frac{a_1}{-a_2}} \operatorname{sech}(-\sqrt{a_1 a_3} \ell t) \\
 m_2 &= \pm \ell \tanh(-\sqrt{a_1 a_3} \ell t) \\
 m_3 &= \pm \ell \sqrt{\frac{a_3}{-a_2}} \operatorname{sech}(-\sqrt{a_1 a_3} \ell t) \tag{5.2}
 \end{aligned}$$

$$I = k(\text{constant})$$

$$\theta = \frac{k}{J_1} t + \theta^0$$

To show that transverse homoclinic orbits occur for $\epsilon \neq 0$ we need only show that the Melnikov function

$$M(\theta^0) = \frac{1}{\Omega(I)} \int_{-\infty}^{\infty} \{ \{ F, H^1 \} \} dt, \tag{5.3}$$

has simple zeros, all other conditions of Theorem 4.3 being immediately

satisfied. Note that we must set $k > 0$, so that $I > 0$ for the unperturbed system and hence the inversion of $H^\epsilon = h$ goes through. This is the reason for choosing an attachment with a heavy 'in balance' symmetric component). The KKS bracket is given by (3.6):

$$\begin{aligned} \{ \{ F(m), H^1(m, I, \theta - \theta^0) \} \} &= -m \cdot \nabla_m F \times \nabla_m H^1 \\ &= \left(\frac{a_1}{I_1^2} + \frac{a_2 \cos^2 \theta}{I_2^2} + \frac{a_3 \sin^2 \theta}{I_3^2} \right) m_1 m_2 m_3 \end{aligned} \quad (5.4)$$

Noting that on any homoclinic orbit m_1 and m_3 are even while m_2 is odd, it follows that the constant term $(a_1/I_1^2)m_1 m_2 m_3$ vanishes in the Melnikov integral and we are left with

$$\begin{aligned} M(\theta^0) &= \frac{1}{(k/J_1)} \int_{-\infty}^{\infty} m_1 m_2 m_3 \left(\frac{a_2}{I_2^2} \cos^2 \theta + \frac{a_3}{I_3^2} \sin^2 \theta \right) d\theta, \\ &= \frac{J_1}{2k} \int_{-\infty}^{\infty} m_1 m_2 m_3 \left(\frac{a_2}{I_2^2} \cos 2\theta - \frac{a_3}{I_3^2} \cos 2\theta \right) d\theta, \\ &= \frac{J_1}{2k} \left(\frac{a_2}{I_2^2} - \frac{a_3}{I_3^2} \right) \int_{-\infty}^{\infty} m_1 m_2 m_3 \cos 2\theta d\theta. \end{aligned} \quad (5.5)$$

Inserting the expressions for the homoclinic orbits, we get

$$\begin{aligned} M(\theta^0) &= \frac{J_1}{2k} \left(\frac{a_2}{I_2^2} - \frac{a_3}{I_3^2} \right) \int_{-\infty}^{\infty} \left(\frac{\ell^3 \sqrt{a_1 a_3}}{-a_2} \right) \operatorname{sech}^2(-\sqrt{a_1 a_3} \ell t) \tanh(-\sqrt{a_1 a_3} \ell t) \\ &\quad \cos 2 \left(\frac{k}{J_1} t + \theta^0 \right) dt, \\ &= C \left[\int_{-\infty}^{\infty} \operatorname{sech}^2(-\sqrt{a_1 a_3} \ell t) \tanh(-\sqrt{a_1 a_3} \ell t) \sin \left(\frac{2kt}{J_1} \right) dt \right] \sin 2\theta^0, \end{aligned} \quad (5.6)$$

where

$$C = \frac{j_1}{2k} \left(\frac{a_2}{I_2^2} - \frac{a_3}{I_3^2} \right) \frac{\ell^3 \sqrt{a_1 a_3}}{a_2} .$$

The integral of (5.6) may be evaluated by the method of residues to obtain

$$M(\theta^0) = C \cdot \frac{2\pi k^2}{j_1^2 (-\sqrt{a_1 a_3} \ell)^3} \operatorname{cosech} \left(\frac{k\pi}{-\sqrt{a_1 a_3} \ell j_1} \right) \sin 2\theta^0 , \quad (5.7)$$

which has simple zeros. We therefore have

5.1 Theorem. The free rigid body with a single slightly asymmetrical freely rotating attachment on one of its principal axes possesses transverse heteroclinic orbits and hence Smale horseshoes in a suitably chosen cross section of the constant energy surface with $k > 0$.

This implies that the rigid body equations with an additional attachment are non integrable. More precisely, if we make a Markov partition of the invariant sphere consisting of the four open regions filled with periodic notions in the unperturbed case (Figure 1), then the dynamics of the perturbed Poincaré map is conjugate to the subshift of finite type on these four symbols. To see this we sketch the homoclinic structure on the sphere in Figure 5, identifying one of the centers (in region A) with the point at infinity. It is clear that orbits starting near the manifold on the 'boundary' of regions 2 and 3 can be selected such that they pass either from region $2 \rightarrow 2$ or $2 \rightarrow 3$ or $3 \rightarrow 2$ or $3 \rightarrow 3$. Similarly on the border of 1,2 orbits can be found passing from $1 \rightarrow 1$, $1 \rightarrow 2$, $2 \rightarrow 1$ or $2 \rightarrow 2$. Continuing in this

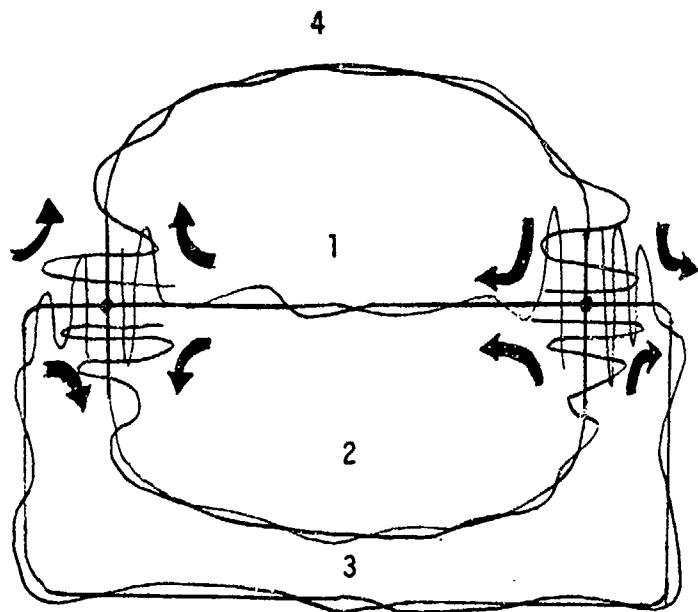


Figure 5. The homoclinic structure of the perturbed manifold $W^s(0, \pm\ell, 0)$, $W^n(0, \pm\ell, 0)$ on the sphere.

way we find the transition matrix

$$A = [a_{ij}] = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}$$

where $a_{ij} = 1$ if there is an orbit from region i to region j and $a_{ij} = 0$ if there is no such orbit.

If two or more attachments are added to the free rigid body then Arnold diffusion may take place. Taking a system with two identical attachments on axis 1, each free to rotate independently, with momenta I, J and angles θ, ϕ , we have the Hamiltonian

$$\begin{aligned}
 H^\epsilon = F + G_1 + G_2 + \epsilon H^1 + O(\epsilon^2) &= \frac{1}{2} \left[\sum_{j=1}^3 \frac{m_j^2}{I_j} + \frac{I^2 + J^2}{J_1} \right] \\
 &- \frac{\epsilon}{2} \left\{ \frac{m_1^2}{I_1^2} + \frac{m_2^2}{I_2^2} (\cos^2 \theta + \cos^2 \phi) + \frac{m_3^2}{I_3^2} (\sin^2 \theta + \sin^2 \phi) + \frac{I^2 + J^2}{J_1^2} \right\} + O(\epsilon^2)
 \end{aligned} \tag{5.8}$$

In this case the Melnikov vector consists of the pair of function

$$\begin{aligned}
 M_1(\theta^0, \phi^0) &= \int_{-\infty}^{\infty} \{I, H^1\} dt \\
 M_2(\theta^0, \phi^0) &= \int_{-\infty}^{\infty} \frac{1}{\Omega_2(k)} \{\{F, H^1\}\},
 \end{aligned}$$

where $\{I, H^1\} = -\frac{\partial H^1}{\partial \theta}$ is the usual canonical Poisson bracket and $\Omega_2 = \frac{J}{J_2}$.

Computations similar to those above (also cf. Holmes and Marsden [1981c], §4) show that

$$\begin{aligned}
 M_1 &= C_1 \sin 2\theta_1^0 \\
 M_2 &= C_2 \sin 2\theta_1^0 + C_3 \sin 2\theta_2^0
 \end{aligned} \tag{5.10}$$

where the C_j are non-zero constants depending on I, J, J_1 , and the choice of total energy and energy in each 'mode' ($E = \ell^2/2I_2 + k_1^2/2J_1 + k_2^2/2J_1$).

Thus M has simple (transverse) zeros for $\theta^0 = m\pi, \phi^0 = n\pi, m, n \in \mathbb{Z}$.

Since $\Omega_1'(I) = \frac{1}{J_1} \neq 0$, all the conditions (H1) - (H4) of Theorem 4.4 hold.

Therefore, the rigid body with two attachments as specified exhibits Arnold diffusion. The existence of a transition chain of two-tori connected by

heteroclinic orbits, (see Holmes and Marsden [1981c], §3 for a discussion) implies that angular momentum can be transferred back and forth between the two spinning attachments in a chaotic manner.

Remark. An amusing corollary for the case of a single attachment is that there are configurations of the Euler elastica for which the sequence of loops above and below the mean level can be prescribed in advance. (For example the loops can be coded by the binary expansion of an irrational number). This follows from the above calculations and the remarkable fact that the elastica equations have the form of the equations of a rigid body with an attachment; see Love [1927, p. 400].

§6. Melnikov Theory for Systems with S^1 Symmetry

We now develop a version of the Melnikov theory that applies to perturbations of a two degree of freedom system with an S^1 symmetry. We have chosen this context with applications to the motion of a nearly symmetric heavy top in mind.

The key new feature is that the unperturbed system is no longer assumed to be a product system consisting of variables with a homoclinic orbit and action angle variables. Rather, this product structure is generalized to the assumption of an S^1 reduction.

Roughly speaking, our unperturbed Hamiltonian H^0 no longer can be split as

$$H^0(q, p, I) = F(q, p) + G(I)$$

so that the frequency function $\Omega = \frac{\partial H^0}{\partial I}$ now may depend on (p, q) . This is, in fact the situation for the nearly symmetric heavy top.

Let us start with a four dimensional symplectic manifold P , whose points are denoted x . Suppose S^1 acts on P by canonical transformations and has an Ad^* -equivariant momentum map $J : P \rightarrow \mathbb{R}$. Let the reduced space be denoted

$$P_\mu = J^{-1}(\mu)/S^1$$

(see Marsden and Weinstein [1974]).

For the heavy top, $P = T^*S^2$ and S^1 consists of rotations about the axis of symmetry. To keep the notation consistent, we shall use ψ for the angle on S^1 . Motivated by this example, we allow P_μ to have isolated singularities, but in this case we demand that the constructions carried out below make sense at the singular points. For the heavy top this causes no difficulties.

Let points in the reduced space be denoted $u \in P_\mu$. Thus, u consists of an S^1 orbit in P , lying in the level set $J^{-1}(\mu)$. Choosing a slice (cross-section) for this action, ψ parametrizes the point on the orbit and of course the value of J labels the surface $J^{-1}(\mu)$. Thus, we write points x as

$$x = (u, \psi, J) \quad (6.1)$$

See Figure 6.

In this notation, a function of x is S^1 invariant if and only if it can be expressed as a function of u and J alone. Such functions induce functions on P_μ , the reduced space, and have well defined Poisson brackets on P_μ . In fact, the Poisson bracket of two such functions is just that function induced on P_μ by their Poisson bracket on P . Of course any S^1 invariant Hamiltonian on P gives a completely integrable system, the integrals being J and H , or equivalently H for the reduced one degree of freedom system.

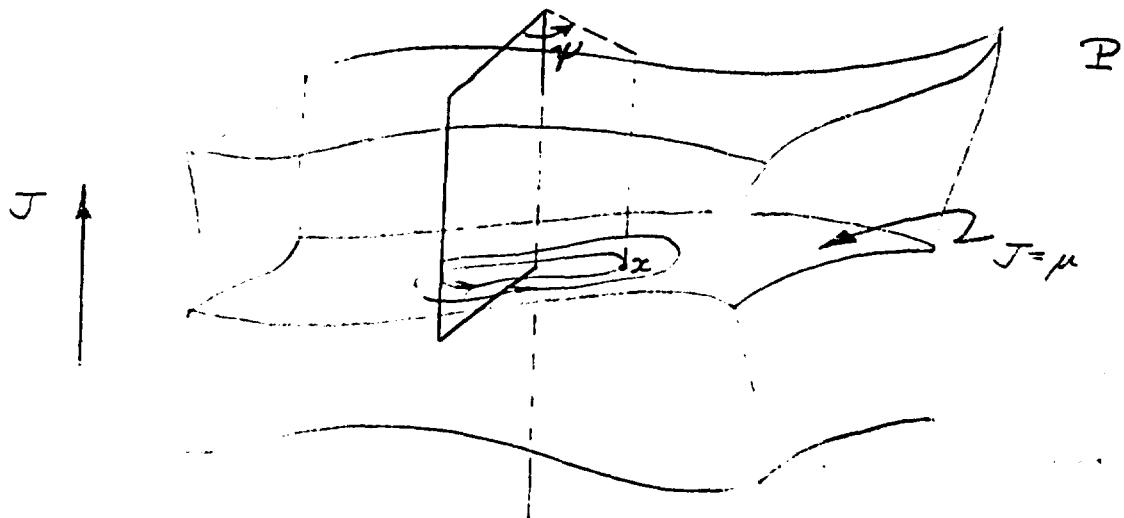


Figure 6

Now assume that we start with an S^1 invariant Hamiltonian $H^0(u, J)$ and we perturb it by a general Hamiltonian. Thus, write

$$H^\varepsilon(u, \psi, J) = H^0(u, J) + \varepsilon H^1(u, \psi, J) + O(\varepsilon^2) \quad (6.2)$$

Now we must generalize the procedure of Holmes and Marsden [1981b] to this context.

We begin by letting

$$\Omega(u, J) = \frac{\partial H^0}{\partial J}(u, J) \quad (6.3)$$

and assuming Ω is positive, so that $H^0(u, J)$ is invertible in the J -variable. For ε small then

$$H^\varepsilon(u, \psi, J) = h \quad (6.4)$$

can be solved for J . Write $H_u^0(J) = H^0(u, J)$ and $(H_u^0)^{-1}$ for its inverse in the J variable.

6.1 Proposition. $J = L^\varepsilon(u, \psi, h) = L^0(u, h) + \varepsilon L^1(u, \psi, h) + O(\varepsilon^2)$ where

$$L^0(u, h) = (H_u^0)^{-1}(h) \quad (6.5)$$

and

$$L^1(u, \psi, h) = \frac{-H^1(u, \psi, L^0(u, h))}{\Omega(u, L^0(u, h))} \quad (6.6)$$

Proof. (6.4) reads

$$H^\varepsilon(u, \psi, L^0 + \varepsilon L^1 + O(\varepsilon^2)) = h$$

i.e.

$$H^0(u, L^0 + \varepsilon L^1 + O(\varepsilon^2)) + \varepsilon H^1(u, \psi, L^0 + \varepsilon L^1 + O(\varepsilon^2)) = h$$

i.e. $H^0(u, L^0) + \epsilon \Omega(u, L^0) + \epsilon H^1(u, \psi, L^0) = h + O(\epsilon^2)$. Comparing powers of ϵ gives the result. ■

As in Marsden [1981, Lecture 4] we can arrange things so that ψ and J are conjugate variables. Thus under the dynamics of H^ϵ ,

$$\dot{\psi} = \{ \psi, H^\epsilon \} = \frac{\partial H^\epsilon}{\partial J} \quad (6.7)$$

Now we change time variables from t to ψ and write $'$ for $d/d\psi$. Thus, for a function $F(u)$,

$$F' = \frac{\dot{F}}{\dot{\psi}} = \{ F, H^\epsilon \} / \frac{\partial H^\epsilon}{\partial J} \quad (6.8)$$

However, from (6.4) and properties of the Poisson bracket,

$$\{ F(u), H^\epsilon(u, \psi, L^\epsilon(u, \psi, h)) \} = 0$$

so

$$\{ F, H^\epsilon \} + \frac{\partial H^\epsilon}{\partial J} \{ F, L^\epsilon \} = 0 \quad (6.9)$$

Comparing (6.8) and (6.9),

$$F' = - \{ F, L^\epsilon \} = - \{ F, L^0 \} - \epsilon \{ F, L^1 \} + O(\epsilon^2) \quad (6.10)$$

Thus, (6.10) is in the form of a periodically forced Hamiltonian system on the reduced phase space. The brackets in (6.10) are taken in the u -variable alone. Thus, we have:

6.2 Proposition. The evolution of u as a function of ψ is a ψ -dependent Hamiltonian system in the reduced phase space. The equations of evolution are given by (6.10).

Now suppose that the reduced system for $\epsilon = 0$ has a homoclinic orbit. To detect the presence of transverse homoclinic orbits for $\epsilon \neq 0$ we must

integrate $\{L^0, L^1\}$ around this orbit. Now in examples, L^0 and L^1 and their bracket can be laborious to compute. It will save some effort if we relate $\{L^0, L^1\}$ to H^0 and H^1 .

$$6.3 \text{ Proposition. } \{L^0, L^1\} = \frac{1}{\Omega} \{H^0, \frac{H^1}{\Omega}\} \quad (6.11)$$

where Ω is evaluated at u, J and the brackets are taken with respect to u , holding ψ, J fixed.

Remarks 1. Formula (6.11) is a generalization of (3.1) of Holmes and Marsden [1981b]. In the later case Ω was independent of u .

2. The above development does not depend on the dimension of P .

Proof of 6.3. Given $F(u)$, we first compute $\{L^0, F\}$. From (6.5), $H^0(u, L^0(u, h)) = h$, so

$$\{H^0, F\} + \Omega \{L^0, F\} = 0$$

Thus

$$\{L^0, F\} = -\frac{1}{\Omega} \{H^0, F\} \quad (6.12)$$

Next, from (6.6) we obtain

$$\{G, L^1\} = -\{G, \frac{H^1}{\Omega}\} - \frac{1}{\Omega} \frac{\partial H^1}{\partial J} \{G, L^0\} + \frac{H^1}{\Omega^2} \frac{\partial \Omega}{\partial J} \{G, L^0\} \quad (6.13)$$

Taking $G = L^0$ in (6.13) the last two terms vanish, so

$$\{L^0, L^1\} = -\{L^0, \frac{H^1}{\Omega}\}$$

By (6.12) this gives the stated result. ■

Assembling these results as we did in Section 4, we obtain the following.

6.4 Theorem. Consider a two degree of freedom Hamiltonian system on a symplectic manifold P with an S^1 symmetry with a Hamiltonian of the form (6.2). Assume that J is chosen and fixed so that the reduced system $H^0(u, J)$ has a homoclinic (or heteroclinic) orbit $\bar{u}(t)$ in the reduced space P_μ and so $\Omega(\bar{u}(t), J) > 0$. Let

$$\psi(t) = \int_0^t \Omega(\bar{u}(t), J) dt + \psi^0 \quad (6.14)$$

and let $\{H^0, \frac{H^1}{\Omega}\} (t, \psi^0)$ denote the u -Poisson bracket evaluated at $\bar{u}(t)$, $\psi(t)$ and J . Let

$$M(\psi^0) = \int_{-\infty}^{\infty} \{H^0, \frac{H^1}{\Omega}\} (t, \psi^0) dt \quad (6.15)$$

and assume $M(\psi^0)$ has simple zeros as a function of ψ^0 . Then for ϵ sufficiently small, the system (6.2) has transverse homoclinic (or heteroclinic) orbits, and hence Smale horseshoes on the energy surface $H^\epsilon = h$, where $h = H^0(\bar{u}, J)$.

§7. Example: the motion of a nearly symmetric heavy top

We now show that Theorem 6.4 implies:

7.1 Theorem. If I_1/I_3 is sufficiently large, $I_2 = I_1 + \epsilon$ and ϵ is sufficiently small, $\epsilon \neq 0$, then the Hamiltonian system for heavy top (see 3.11 and 3.12) has transverse homoclinic orbits in the Poincaré map for the ψ variable on energy surfaces close to the homoclinic orbit described in 3.2 or 3.4.

7.2 Corollary. The heavy top close to the symmetric top has no analytic integrals other than the energy and angular momentum about the vertical axis.

Remarks 1. As we have already discussed, this corollary has recently been obtained by Ziglin, but by rather different methods. Moreover, our result 7.1 shows the existence of 'chaotic' orbits.

2. I_1/I_3 being large can be replaced by $0 < \gamma < 1$ and the integral 7.13 below being non-zero. This integral is non-zero for most values of M, g, I_1, I_3 , as we shall show.

We shall prove theorem 7.1 in the Euler angle representation first and then sketch how the proof can be alternatively obtained using the KKS description.

In (3.12) let $I_2 = I_1 + \epsilon$. This gives

$$H^0 = \frac{1}{2} \left\{ \frac{(p_\phi - p_\psi \cos \theta)^2}{I_1 \sin^2 \theta} + \frac{p_\theta^2}{I_1} + \frac{p_\psi^2}{I_3} \right\} + Mg\ell \cos \theta \quad (7.1)$$

and

$$H^1 = -\frac{1}{2I_1^2 \sin^2 \theta} ((p_\phi - p_\psi \cos \theta) \cos \psi - p_\theta \sin \theta \sin \psi)^2 \quad (7.2)$$

Note that H^0 is the Hamiltonian for the symmetric top and so has a homoclinic orbit given by (3.15).

Since ψ and p_ψ are conjugate variables and $J = p_\psi$, we have from (6.3),

$$\Omega = \frac{p_\psi}{I_3} - \left(\frac{p_\phi - p_\psi \cos \theta}{I_1 \sin^2 \theta} \right) \cdot \cos \theta \quad (7.3)$$

On the homoclinic orbit, $p_\phi = p_\psi = \text{constant} \equiv I_1 b$, so (7.3) becomes

$$\Omega = b \left(\frac{I_1}{I_3} - \frac{\cos \theta}{1 + \cos \theta} \right) \quad (7.4)$$

7.3 Lemma. With $p_\phi = p_\psi = I_1 b$, $\beta = \frac{2 Mg\ell}{I_1}$,

$$\dot{\phi} = \frac{b}{1 + \cos \theta} \quad \text{and} \quad \dot{\theta} = \frac{p_\theta}{I_1} \quad ,$$

we have

$$\{H^0, \frac{H^1}{\Omega}\}_{\theta, p_\theta} = \frac{1}{\Omega} (A(\theta) \sin 2\psi + B(\theta) \cos 2\psi + C(\theta)) \quad (7.5)$$

where

$$A(\theta) = \frac{b}{2} \left(\dot{\phi}^2 - \frac{\beta}{2} + \frac{\dot{\theta}^2 \phi^2}{b\Omega} \right) (1 - \cos \theta) - \dot{\theta}^2$$

$$B(\theta) = \left(\dot{\theta} \dot{\phi}^2 - \frac{\beta \dot{\theta}}{4} + \frac{\dot{\theta}^3 \phi^2}{4b\Omega} \right) \sin \theta - \frac{\dot{\theta} \dot{\phi}^4}{4b\Omega} - \sin^3 \theta$$

and

$$C(\theta) = \dot{\theta} \left(\frac{\dot{\phi}^2}{4} - \frac{\dot{\theta}^3 \dot{\phi}^2 + \dot{\theta}^2 \phi^2}{4b\Omega} \right) \sin \theta$$

Proof. The computations are slightly tedious but straightforward. We write

$$\left\{ H^0, \frac{H^1}{\Omega} \right\} = \frac{1}{\Omega} \{ H^0, H^1 \} - \frac{H^1}{\Omega} \{ H^0, \Omega \} , \quad (7.6)$$

$$\{ H^0, H^1 \} = \frac{\partial H^0}{\partial \theta} \frac{\partial H^1}{\partial p_\theta} - \frac{\partial H^0}{\partial p_\theta} \frac{\partial H^1}{\partial \theta} ,$$

$$\{ H^0, \Omega \} = - \frac{\partial H^0}{\partial p_\theta} \frac{\partial \Omega}{\partial \theta} ,$$

and compute that

$$\frac{\partial H^0}{\partial \theta} = -I_1 \left(\dot{\phi}^2 \sin \theta - \frac{\beta}{2} \sin \theta \right) ,$$

$$\frac{\partial H^0}{\partial p_\theta} = \dot{\theta}$$

$$\frac{\partial H^1}{\partial \theta} = -\dot{\phi}^2 \sin \theta \cos^2 \psi + \frac{\dot{\phi} \dot{\theta}}{2} \sin 2\psi ,$$

$$\frac{\partial H^1}{\partial p_\theta} = \frac{1}{I_1} \left(\frac{\dot{\phi}}{2} \sin \theta \sin 2\psi - \dot{\theta} \sin^2 \psi \right) ,$$

$$\begin{aligned} \{ H^0, H^1 \} &= \frac{\dot{\phi}}{2} \left[\left(\dot{\phi}^2 - \frac{\beta}{2} \right) \sin^2 \theta - \dot{\theta}^2 \right] \sin 2\psi \\ &\quad + \dot{\theta} \dot{\phi}^2 \sin \theta \cos 2\psi - \frac{\beta \dot{\theta}}{4} \sin \theta (\cos 2\psi - 1) , \end{aligned}$$

$$\frac{\partial \Omega}{\partial \theta} = \frac{\dot{\phi}^2}{b} \sin \theta ,$$

and

$$\{ H^0, \Omega \} = - \frac{\dot{\theta} \dot{\phi}^2}{b} \sin \theta .$$

Substituting these expressions into (7.6) and simplifying yields (7.5). ■

The Melnikov function is given by

$$M(\psi^0) = \int_{-\infty}^{\infty} \left\{ H^0, \frac{H^1}{\Omega} \right\} dt \quad (7.7)$$

where the integral is evaluated along the homoclinic orbit

$$\cos \theta = 1 - \gamma \operatorname{sech}^2 \left(\frac{\sqrt{B\gamma}}{2} t \right), \quad p_\theta = I_1 \dot{\theta} \quad (7.8)$$

(see (3.15)) and where

$$\psi(t) = \int_0^t \Omega(t) dt + \psi^0 = \bar{\psi}(t) + \psi^0 \quad (7.9)$$

and

$$\Omega(t) = b \left(\frac{I_1}{I_3} + \frac{1 - \gamma \operatorname{sech}^2 \left(\frac{\sqrt{B\gamma}}{2} t \right)}{\gamma \operatorname{sech}^2 \left(\frac{\sqrt{B\gamma}}{2} t \right)} \right) \quad (7.10)$$

Note that if $0 < \gamma < 1$, or if I_1/I_3 is sufficiently large, $\Omega(t) > 0$.

Substitute (7.5) and (7.9) into (7.7) yields

$$\begin{aligned} M(\psi^0) &= \left[\int_{-\infty}^{\infty} \frac{1}{\Omega} (A(\theta) \sin 2\bar{\psi} + B(\theta) \cos 2\bar{\psi}) dt \right] \cos 2\psi^0 \\ &+ \left[\int_{-\infty}^{\infty} \frac{1}{\Omega} (A(\theta) \cos 2\bar{\psi} - B(\theta) \sin 2\bar{\psi}) dt \right] \sin 2\psi^0 + \int_{-\infty}^{\infty} \frac{1}{\Omega} C(\theta) dt \end{aligned} \quad (7.11)$$

The first, second and fifth terms are odd functions of t and so we obtain the following

7.4 Lemma.

$$M(\psi^0) = \left[\int_{-\infty}^{\infty} \frac{1}{\Omega} (A(\theta) \cos 2\bar{\psi} - B(\theta) \sin 2\bar{\psi}) dt \right] \sin 2\psi^0 \quad (7.12)$$

Now we observe that at $t = 0$, $A = \frac{b\gamma^2}{4(2-\gamma)} > 0$ and $B = 0$. It follows that for I_1/I_3 sufficiently large, the portion of the integral from the first term near $t = 0$ dominates and so the integral must be a non-zero number. Thus $M(\psi^0)$ has simple zeros and the theorem is proved.

Remark. For any given b, β, I_1, I_3 (with $\gamma = 2 - \frac{b^2}{\beta}$ and $b^2 < 2\beta$, $b, \beta > 0$) one needs only the condition that

$$\int_{-\infty}^{\infty} \frac{1}{\Omega} (A(\theta) \cos 2\bar{\psi} - B(\theta) \sin 2\bar{\psi}) dt \quad (7.13)$$

be non-zero. Since we cannot evaluate (7.13) analytically it does not seem so simple to decide exactly when (7.13) vanishes. Since we can be sure it is non-zero for I_1/I_3 large, it follows that (7.13) can vanish for at most a finite set of values of b, β, I_1, I_3 as it is analytic. Thus, we can be sure of transverse homoclinic orbits for generic M, I_1, I_3, b , if $0 < \gamma < 1$.

Finally, we indicate how the same computations can be done using the KKS formalism. This actually makes the computations slightly easier, but the final result is the same.

Again, letting $I_2 = I_1 + \varepsilon$ in $H = \frac{1}{2} \sum_{j=1}^3 \frac{m_j^2}{I_j} + Mg\ell v_3$, we have

$$H^0 = \frac{m_1^2 + m_2^2}{2I_1} + \frac{m_3^2}{2I_3} + Mg\ell v_3 \quad (7.14)$$

and

$$H^1 = -\frac{m_2^2}{2I_1^2}. \quad (7.15)$$

Next, observe that ψ and m_3 are conjugate variables, where $\tan \psi = v_1/v_2$, so if $m \cdot v = I_1 b = m_3$,

$$\Omega = \{\{\psi, H^0\}\} = b \left(\frac{I_1}{I_3} - \frac{v_3}{1 + v_3} \right) \quad (7.16)$$

which agrees with (7.4). To compute $\{\{H^0, \frac{H^1}{\Omega}\}\}$ we write (see Lemma 3.3) H^1 as a function of the reduced variables $v_2 m_1 - v_1 m_2$ and v_3 and compute the bracket holding ψ and m_3 fixed. This is done by writing

$$\begin{aligned} m_2^2 &= \frac{1}{1 - v_3^2} (m \cdot v - m_3 v_3) \cos^2 \psi - (v_2 m_1 - v_1 m_2) (m \cdot v - m_3 v_3) \sin \psi \\ &\quad + (v_2 m_1 - v_1 m_2)^2 \sin^2 \psi \end{aligned} \quad (7.17)$$

When $m \cdot v = m_3 = I_1 b$, this becomes

$$m_2^2 = I_1 b \frac{\cos^2}{1 + v_3} - I_1 b (1 - v_3) (v_2 m_1 - v_1 m_2) \sin \psi + (v_2 m_1 - v_1 m_2)^2 \sin^2 \psi \quad (7.18)$$

Using these expressions we readily compute $\{\{H^0, \frac{H^1}{\Omega}\}\}$ using (3.17). After substituting from Table 1, the same expression (7.11) as above results. Thus, the proof may be completed in the same way as with the Euler angle argument.

References

R. Abraham and J. Marsden [1978]. Foundations of Mechanics, second edition, Addison-Wesley, Reading, MA.

L.D. Akulenko, D.D. Leshchenko and F.L. Chernousko [1979]. Perturbed motions of a rigid body close to the Lagrange case.

J. Arms, J. Marsden and V. Moncrief [1981], Symmetry and bifurcations of momentum mappings. Comm. Math. Phys., 78, pp. 455-478.

V. I. Arnold [1964]. Instability of dynamical systems with several degrees of freedom, Dokl. Akad. Nauk. SSSR 156, 9-12.

V. I. Arnold [1966]. Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. Ann. Inst. Fourier Grenoble 16, 319-361.

V. Arnold [1978]. Mathematical methods of classical mechanics. Springer Graduate Texts in Math. No. 60 Springer-Verlag, New York.

D. Ebin and J. Marsden [1970]. Groups of diffeomorphisms and the motion of an incompressible fluid, Ann. of Math. 92, 102-163.

H. Goldstein [1980]. Classical Mechanics, 2nd Ed. Addison-Wesley.

B. Greenspan and P.J. Holmes [1981]. Homoclinic Orbits, Subharmonics and Global Bifurcations in Forced Oscillations. In 'Non-linear Dynamics and Turbulence' ed. G. Barenblatt, G. Iooss and D.D. Joseph, Pitman (to appear).

V. Guillemin and S. Sternberg [1980]. The moment map and collective motion, Ann. of Phys. 127, 220-253.

P. Holmes [1980]. Averaging and chaotic motions in forced oscillations, Siam. J. on Appl. Math. 38, 65-80 and 40, 167-168.

P.J. Holmes and J.E. Marsden [1981a]. A partial differential equation with infinitely many periodic orbits: chaotic oscillations of a forced beam, Arch. Rat. Mech. Anal. (to appear).

P.J. Holmes and J.E. Marsden [1981b]. Horseshoes in perturbations of Hamiltonian systems with two degrees of freedom. Comm. Math. Phys. (to appear).

P.J. Holmes and J.E. Marsden [1981c]. Melnikov's method and Arnold diffusion for perturbations of integrable Hamiltonian systems (preprint).

A.E.H. Love [1927]. A Treatise on the Mathematical Theory of Elasticity, Cambridge University Press (Reprinted by Dover [1944]).

J. Marsden [1981]. Lectures on Geometric Methods in Mathematical Physics, CBMS-NSF Regional Conference Series #37, SIAM.

J. Marsden and A. Weinstein [1974]. Reduction of symplectic manifolds with symmetry, Rep. Math. Phys., 5, pp. 121-130.

J. Marsden and A. Weinstein [1981]. The Hamiltonian structure of the Maxwell-Vlasov equations, Physica D. (to appear).

V.K. Melnikov [1963]. On the stability of the center for time period. perturbations, Trans. Moscow Math. Soc. 12, pp. 1-57.

T. Ratiu and P. van Moerbeke [1981]. The Lagrange rigid body motion (preprint).

S.L. Ziglin [1980]. Decomposition of separatrices, branching of solutions and nonexistence of an integral in the dynamics of a rigid body, Trans. Moscow Math. Soc. 41, 287.

S.L. Ziglin [1981]. Branching of solutions and nonexistence of integrals in Hamiltonian systems. Doklady Akad. Nauk. SSSR 257, 26-29.